# ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ УЧРЕЖДЕНИЕ ФЕДЕРАЛЬНЫЙ НАУЧНЫЙ ЦЕНТР ИССЛЕДОВАНИЙ И РАЗРАБОТКИ ИММУНОБИОЛОГИЧЕСКИХ ПРЕПАРАТОВ ИМ. М.П. ЧУМАКОВА РАН

На правах рукописи

Простова Мария Андреевна

## Структурно-функциональная характеристика апикального участка домена d репликативного элемента oriL генома полиовируса

Диссертация на соискание ученой степени кандидата биологических наук 03.02.02 – вирусология Научный руководитель: к.б.н. Гмыль А.П.

MOCKBA - 2017

Посвящаю своим маме и папе, которые не перестают учиться новому.

## Оглавление

Ог.	Оглавление				
Сп	исок сокращен	ий	4		
Вв	едение		6		
1.	Обзор литер	атуры	12		
	1.1 O	бщие сведения о вирусе полиомиелита	12		
	1.1 Ці	икл репродукции вируса полиомиелита	13		
	1.1.1	Проникновение вирусной РНК в клетку	15		
	1.1.2	Трансляция вирусной РНК и процессинг полипротеина	15		
	1.1.3	Сопряжение трансляции и репликации	17		
	1.1.4	Репликация вирусного генома	19		
	1.1.5	Сборка капсида и упаковка дочерних молекул (+)РНК	26		
	1.2 Ba	заимное узнавание 3CD и oriL: структурно-функциональная характер	оистика27		
	1.2.1	Белок 3CD	27		
	1.2.2	Репликативный элемент oriL: характеристика домена d			
	1.3 П <sub>1</sub>	ространственные структурные классы тетрапетель			
	1.3.1	Пространственная структура тетрапетель класса UNCG			
	1.3.2	Пространственная структура тетрапетель класса GNRA			
	1.3.3	Пространственная структура тетрапетель класса gCUUGc	40		
2.	Материалы и	а методы	41		
	2.1 Pe	агенты	41		
	2.2 Ш	таммы E.coli и получение компетентных клеток	42		
	2.2.1	Штамм ТОР10	42		
	2.2.2	Штамм ЈМ109	43		
	2.2.3	Получение компетентных клеток E.coli	43		
	2.3 Пл	пазмиды и базовые генно-инженерные методы	44		
	2.3.1	Плазмида pT7PV1(Rib+)MS	44		

1

	2.3.2	Конструирование плазмиды pT7PV1Rib(+)MS с рандомизированным	
	октан	уклеотидом в домене d	
	2.3.3	Конструирование плазмид pT7PV1Rib(+)MS с измененными oriL	
	2.3.4	Плазмида pQE60	
	2.3.5	Конструирование плазмиды, экспрессирующей вирусный белок 3CD	
	2.3.6	Трансформация компетентной культуры клеток плазмидной ДНК	
	2.3.7	Выделение плазмидной ДНК	
	2.3.8	Обработка плазмидной ДНК эндонуклеазами	
	2.3.9	Лигирование ДНК	
	2.3.10	Аналитический электрофорез в агарозном геле	
	2.4 Tp	анскрипция и очистка РНК	
	2.4.1	Транскрипция in vitro	
	2.4.2	Очистка полногеномных транскриптов в сахарозном градиенте	
	2.4.3	Получение и очистка oriL	
	2.5 Ку	льтуры клеток и вирусологические методы	
	2.5.1	Культуры перевиваемых клеток и их ведение	
	2.5.2	Метод бляшек	
	2.5.3	Определение титра вируса	
	2.5.4	Размножение вируса	
	2.5.5	Трансфекция культуры эукариотических клеток	
	2.6 Or	пределение последовательности вирусного генома и кинетики синтеза в	зирусной
РНК	53		
	2.6.1	Выделение тотальной РНК из инфицированной культуры клеток	53
	2.6.2	Выделение вирусной РНК	
	2.6.3	Обратная транскрипция in vitro	
	2.6.4	Проведение ПЦР и выделение фрагментов ДНК из геля	
	2.6.5	Кинетика синтеза вирусной РНК	
	2.7 По	олучение рекомбинантного вирусного белка 3CD и эксперименты по свя	зыванию
	oriL in v	itro	
	2.7.1	Экспрессия и очистка вирусного белка 3СD	
	2.7.2	Получение антител мыши к рекомбинантному белку 3CD	
		, <u>1</u> , <u>,</u>	2

	2.7.3 Метод EMSA
	2.7.4 Визуализация комплексов 3CD/oriL методом Western-blot
	2.8 Использованное программное обеспечение
3.	Результаты61
	3.1 Отбор жизнеспособных вариантов полиовируса из набора случайных
	3.1.1 Оценка качества олигонуклеотида, использованного для рандомизации61
	3.1.2 Отбор жизнеспособных вариантов полиовируса
	3.1.3 Разнообразие последовательностей в апикальном участке домена d у
	экспериментально отобранных вирусов67
	3.1.4 Разнообразие последовательностей тетрапетли домена d и фланкирующей пары
	у вирусов вида Enterovirus C70
	3.2 Конструирование полиовирусов, содержащих мутации в апикальном участке домена
	d, и эффективность их репродукции75
	3.3 Эффективность синтеза дочерних копий РНК у сконструированных вариантов88
	3.4 Влияние пространственной структуры апикального участка домена d oriL на
	эффективность его взаимодействия с белком 3CD90
4.	Обсуждение
5.	Заключение
6.	Выводы
7.	Список литературы

## Список сокращений

- 5'НТО и 3'НТО 5'- и 3'-нетранслируемые области
- ALV avian leukosis virus, вирус лейкоза птиц
- CD155 cluster of differentiation 155, трансмембранный гликопротеин эукариотической клетки, рецептор для полиовируса (PVR)
- CRE cis-acting replication element, цист-действующий репликативный элемент (oril)
- CVB3, CVB4 Coxackie virus B3, B4, Коксаки вирусы B3, B4
- dsRBD double stranded RNA binding domain, домен, связывающий двуцепочечную РНК
- eIF4G, eIF4E eukaryotic translation initiation factor 4G, 4E, эукариотические факторы инициации трансляции мРНК
- HRV14 human rhinovirus 14, риновирус человека 14
- IRES internal ribosome entry site, внутренний сайт посадки рибосом
- ITAFs IRES translation activation factors, факторы активации IRES-опосредованной трансляции
- КН-домен К-homology домен впервые обнаруженные у ядерного рибонуклеопротеида hnRNP К
- m<sup>7</sup>G 7 метилгуанин, кэп-структура
- mRNA messenger RNA, матричная РНК
- NC nucleocapsid, нуклеокапсид
- NRE nucleolin recognition element, РНК-элемент, узнаваемый нуклеолином
- oriI origine internal, репликативный элемент генома полиовируса, внутренний
- oriL -origin left, репликативный элемент генома полиовируса, левый
- oriR-origin right, репликативный элемент генома полиовируса, правый
- PABP PolyA Binding Protein, белок связывающий последовательность полиА
- PCBP2 PolyC Binding Protein 2, также называемый hnRNP E (heterogeneous nuclear ribonucleoprotein E) рибонуклеопротеид Е формирует один из гетерогенных рибонуклеопротеидов ядра клетки, связывает полиС последовательность

PDB – (Protein Data Bank) база данных

- PDB ID идентификационный код файлов, содержащим координаты и описание молекул, в базе данных Protein Data Bank
- PTB polypyrimidine tract binding protein, белок, связывающий полипиримидиновую последовательность, также формирует один из гетерогенных рибонуклеопротеидов ядра
- RBD RNA binding domain (также называемый RRM или RNP домен)
- Rnt1p рибонуклеаза дрожжей, принадлежит семейству РНКаз III

rRNA – ribosomal RNA, РНК рибосом

RSV - Rous sarcoma virus, вирус саркомы Рауса

SAM - sterile alfa motif, консервативный домен белка Vts1p

SAXS - small angle X-ray scattering, метод мало углового рентгеновского рассеивания

- SELEX (systematic evolution of ligand by exponential enrichment) методический подход, позволяющий отбирать подходящие лиганды с помощью систематической эволюции
- SRE Smaug recognition element, участок РНК, узнаваемый гомологами белка Smaug (?)
- SRp20 SR protein 20 serine/arginine-rich protein splicing factors, богатый серинами и аргининами Белок, участвующий в сплайсинге
- TDP2 tyrosyl-DNA phosphodiesterase 2, тирозил-ДНК фосфодиэстераза 2
- Tm температура плавления
- tRNA transport RNA, транспортная РНК
- VP1 VP2 VP3 VP4 VP0- virus protein 1, 2,3,4,0, структурные вирусные белки
- VPg genome linked virus protein, вирусный белок, связанный с геномом

VPgpUpU<sub>OH</sub> – уридилированный белок VPg

- Vts1p РНК-связывающий белок дрожжей, является гомологом белка Smaug
- ДСН-ПААГ полиакриламидный гель, содержащий додецилсульфат натрия
- РНК (RNA) рибонуклеиновая кислота (Ribonucleic acid)
- ЯМР ядерный магнитный резонанс

## Введение

#### Актуальность темы исследования

Полиовирус является представителем семейства Picornaviridae - безоболочечных вирусов с однонитевым PHK-геномом положительной полярности и небольшим икосаэдрическим капсидом. Инфицирование вирусом полиомиелита может вызывать у человека паралитический полиомиелит. Другие вирусы семейства Picornaviridae вызывают такие заболевания человека и животных, как ящур, ринит, гепатит A, серозный менингит и другие. Регуляция экспрессии и репликации генома пикорнавирусов осуществляется посредством взаимодействия его PHK-элементов с вирусными или клеточными белками. В процессе репликации генома вируса полиомиелита образуется по меньшей три ключевых PHK-белковых комплекса. До сих пор ни для одного из этих комплексов неизвестно, как осуществляется взаимное узнавание его элементов, и какие события следуют за его образованием.

Полимераза энтеровирусов вносит около одной случайной замены на каждый дочерний геном, что с одной стороны, позволяет вирусу быстро адаптироваться к изменяющимся условиям, а с другой стороны ставит проблему сохранения функциональности ключевых элементов генома (Acevedo, Brodsky & Andino, 2014). Исследование пределов изменчивости репликативного элемента oriL позволит приблизиться к пониманию механизмов, лежащих в основе помехоустойчивости вирусов с высокой скоростью изменения генома, и принципов РНК-белкового узнавания.

#### Степень разработанности темы

Известно, что взаимодействие 3CD и *ori*L является необходимым для инициации репликации вирусного генома (Andino, Rieckhof & Baltimore, 1990a; Gamarnik & Andino, 1997, 1998). Известны основные взаимодействующие элементы со стороны обоих партнеров. Узнаваемая белком 3CD тетрапетля находится в апикальном участке домена d репликативного элемента *oriL* (Trono, Andino & Baltimore, 1988; Andino *et al.*, 1990a, 1990b). Показано, что количество нуклеотидов в апикальном участке домена d играет значимую роль для взаимодействия с белком 3CD (Zell *et al.*, 2002). Со стороны белка 3CD известны две аминокислотные последовательности, вовлеченные во взаимодействие с *ori*L ( $_{154}$ TGK $_{156}$  и  $_{82}$ KFRDIR $_{87}$ ), и ряд дополнительных аминокислот (Andino *et al.*, 1990b, 1993; Hämmerle, Molla & Wimmer, 1992). Однако, как осуществляется узнавание вирусного

белка 3CD и oriL остается не ясным.

**Целью** работы является выявление свойств апикального участка домена d репликативного элемента *oriL* генома вируса полиомиелита, необходимых для эффективного взаимодействия с вирусным белком 3CD. Для достижения поставленной цели были сформулированы следующие задачи:

1. установить, какие нуклеотидные последовательности в апикальном участке домена d *oriL* могут поддерживать эффективную репродукцию вируса полиомиелита;

2. сконструировать варианты полиовируса, содержащие различные последовательности в апикальном участке домена d *oriL*, и охарактеризовать их эффективность репродукции и репликации;

создать систему экспрессии, выделения и очистки рекомбинантного вирусного белка
 3CD;

4. охарактеризовать различные варианты *oriL* по их способности взаимодействовать с рекомбинантным белком 3CD.

### Научная новизна. В данной работе впервые

• продемонстрировано, что тетрапетли с разной последовательностью, но принадлежащие пространственному структурному классу UNCG, эффективно узнаются вирусным белком 3CD и поддерживают эффективную репликацию вирусного генома, тогда как тетрапетли пространственных структурных классов GNRA и gCUUGc не могут быть эффективно узнаны этим белком и не могут поддерживать эффективную репликацию

• показано, что вирусы, содержащие различные структурные нарушения апикального участка домена d oriL, даже значительно нарушающие функциональность этого участка, сохраняют жизнеспособность и приобретают мутации, возвращающие последовательность апикального участка домена d к консенсусам nYNHGm (но не gCUUGc) или uGNUAg, или приводящую к аминокислотной замене Thr<sub>154</sub>Ile.

## Теоретическая и практическая значимость работы

В ходе работы было установлено, что для эффективной репликации вируса в апикальном участке домена d репликативного элемента *oriL* необходимо наличие такой пространственной структуры, которая может быть реализована большим количеством различных

7

последовательностей. В результате, при появлении мутаций, этот ключевой элемент сохраняет свою функциональность. Таким образом, для полиовируса принцип узнавания пространственной структуры РНК-элемента, а не его последовательности, может быть дополнительным преимуществом, позволяющим сохранять функциональность генома при значительной скорости появления мутаций.

Результаты данной работы могут быть применены в рациональном дизайне потенциальных ингибиторов вирусной репликации.

### Методология и методы исследования

Для того чтобы определить множество последовательностей в апикальном участке домена d oriL, способных поддерживать жизнеспособность вируса полиомиелита, был осуществлен отбор инфекционных вариантов из набора вариантов со случайной последовательностью. Для этого восемь нуклеотидов апикального участка домена d были рандомизированы, и полученная полногеномная PHK была использована для трансфекции культуры клеток Vero. Жизнеспособные варианты образовывали зоны лизиса культуры клеток под покрытием – бляшки, материал из которых был использован для определения последовательности вирусных геномов. Анализ набора полученных таким образом жизнеспособных вариантов выявил необходимые для поддержания функциональности свойства домена d. Выводы были подтверждены в следующих экспериментах: 1) создание ряда дополнительных мутантов и их характеристика для проверки сформулированных гипотез, 2) характеристика эффективности синтеза дочерних (+) PHK у различных мутантов и 3) исследование эффективности связывания различных мутантных *oriL* с рекомбинантным белком 3CD.

Личный вклад автора. Основная часть экспериментов была осуществлена автором. Эксперимент по отбору жизнеспособных вариантов и анализ полученных результатов был осуществлен сотрудником ИПВЭ им. М.П. Чумакова <u>Бахмутовым Д. В.</u> Автор получил 16 из 33 использованных в работе генно-инженерных конструкций с различной последовательностью в апикальном участке домена d. Эксперименты по определению кинетики синтеза дочерних копий РНК были спланированы автором и осуществлены совместно с сотрудником ИПВЭ им. М.П. Чумакова Шишовой А.А. Разработка метода получения и очистки рекомбинантного белка 3CD была осуществлена автором с использованием рекомендаций сотрудников Института белка РАН Тищенко С.А. и Гарбер М.А. Создание выборки полногеномных последовательностей из базы данных и их выравнивание было осуществлено сотрудником Института А.А. Девяткиным. Анализ выбранных из базы данных последовательностей был осуществлен автором. Получение рекомбинантного белка 3CD и мутантных *oriL* и эксперименты по определению эффективности их взаимодействия были осуществлены автором. Бляшечные фенотипы для половины исследованных мутантов были определены автором совместно с сотрудниками ИПВЭ им. М.П. Чумакова Хитриной Е.В. и Шишовой А.А. Работа с культурами клеток проводилась сотрудниками ИПВЭ им. М.П. Чумакова Хитриной Е.В. и Колесниковой М.С. Иммунизация мышей, их вскрытие, получение сыворотки крови и гипериммунной асцитной жидкости были осуществлены сотрудником ИПВЭ им. М.П. Чумакова Роговой Ю.В. совместно с автором.

## Положения, выносимые на защиту

- эффективная репродукция вируса полиомиелита с природной последовательностью 3С поддерживается множеством тетрапетель в апикальном участке домена d *oriL*, которое можно описать консенсусами: nYNHGn (за исключением gCUUGc) и uGVUAg, где n и m – любые комплементарные нуклеотиды, H – любой нуклеотид, кроме G, V – любой нуклеотид, кроме U.
- 2. тетрапетли в апикальном участке домена d, имеющие пространственную структуру класса UNCG, но разные последовательности, узнаются вирусным белком 3CD и поддерживают эффективную вирусную репликацию. Тетрапетли в апикальном участке домена d, имеющие пространственную структуру классов GNRA и gCUUGc, не узнаются вирусным белком 3CD, и не могут поддерживать эффективную вирусную репликацию. На основании этих данных может быть выдвинута гипотеза об первичном значении пространственной структуры апикального участка домена d oriL для взаимодействия с вирусным белком 3CD.
- 3. аминокислотная замена Thr<sub>154</sub>Ile в белке 3C компенсирует различные нарушения структуры апикального участка домена d и наличие тетрапетель, чья пространственная структура отличается от структуры класса UNCG.

## Степень достоверности и апробация результатов

Достоверность работы обеспечивается достижением сходных результатов с использованием различных методических подходов.

Результаты были представлены на пяти международных конференциях.

- Maria Prostova, Elena Smertina, Denis Bakhmutov, Elena Khitrina, Anatoliy Gmyl, Vadim Agol, Marina Kolesnikova. *Requirements for recognition of the replicative element oriL of poliovirus RNA by the viral protein 3CD*. RNA 2016: The 21st Annual Meeting of the RNA Society, Kyoto, Japan; 07/2016
- Yakovenko M., Gmyl A.P., Prostova M.A., Sherstyuk A., Ivanova O., Eremeeva T., Isaeca O., Agol V.I.. Innocent VDPV? Preliminary results of investigation of vaccine - derived poliovirus with unusual genetic properties. EUROPIC 2012; XVII Meeting of the European Study Group on the Molecular Biology of Picornaviruses, Saint Raphaël, France, 3-7 June 06/2012
- Maria Prostova, Anna Shishova, Ilya Belalov, Backhmutov Denis, Anatoly Gmyl, Elena Khitrina, William Melchers, Jan Zoll, Hans Heus. *Structure-function characterization of poliovirus RNA-element*.
  4th Berlin Summer Meeting Computational and experimental molecular biology "From RNA to Protein and beyond"; 06/2011
- Maria Prostova, Shishova Anna, Backhmutov Denis, Anatoly Gmyl, Elena Khitrina, William Melchers, Jan Zoll, Hans Heus, Vadim Agol: Role of the terminal tetraloop of domain d of replicative cis-acting element oriL of poliovirus genome and its flanking base pairs in oriL-3Cd recognition. EUROPIC 2010; XVIth Meeting in St. Andrews, Scotland; 09/2010
- Maria Prostova, Bakhmutov D.V., Gmyl A.P., Khitrina E.V., Melchers W.J.G., Hues H.A., Agol V.I.: Requirements for optimal interacton of replicative cis-element oriL of poliovirus RNA with its protein ligand, protease 3CD. EUROPIC 2008, XV Meeting of the European Study Group on the Molecular Biology of Picornaviruses, Sitges Barcelona, Spain; 05/2008

## Публикации

По теме диссертации автором опубликовано 3 научных статьи в журналах, индексируемых в международных системах цитирования (Web of Science, Scopus).

- Prostova, M. A., Deviatkin, A. A., Tcelykh, I. O., Lukashev, A. N. and Gmyl, A. P. (2017) 'Independent evolution of tetraloop in enterovirus oriL replicative element and its putative binding partners in protein 3C'. *PeerJ*: 5:e3896, doi: 10.7717/peerj.3896.
- Prostova, M. A., Gmyl, A. P., Bakhmutov, D. V, Shishova, A. A., Khitrina, E. V, Kolesnikova, M. S., Serebryakova, M. V, Isaeva, O. V and Agol, V. I. (2015) 'Mutational robustness and resilience of a replicative cis-element of RNA virus: promiscuity, limitations, relevance.', *RNA biology*, 12(12), pp. 1338–1354. doi:10.1080/15476286.2015.1100794.

Yakovenko, M. L., Gmyl, A. P., Ivanova, O. E., Eremeeva, T. P., Ivanov, A. P., Prostova, M. A., Baykova, O. Y., Isaeva, O. V., Lipskaya, G. Y., Shakaryan, A. K., Kew, O. M., Deshpande, J. M. and Agol, V. I. (2014) 'The 2010 outbreak of poliomyelitis in Tajikistan: epidemiology and lessons learnt.', *Euro surveillance : European communicable disease bulletin*, 19(7), p. 20706. doi:10.2807/1560-7917.ES2014.19.7.20706.

## Структура и объем диссертации

Диссертация представлена на 114 страницах и состоит из оглавления, списка сокращений, введения, обзора литературы, материалов и методов, описания результатов, обсуждения, заключения и списка литературы. В диссертации приведены 32 рисунка и 13 таблиц. Список литературы содержит 246 источников.

## 1. Обзор литературы

## 1.1 Общие сведения о вирусе полиомиелита

Вирус полиомиелита (вид *Human enterovirus C*) принадлежит семейству Picornaviridae, роду *Enterovirus*. Капсид вируса небольшой (27-30 нм), икосаэдрический, без липопротеидной оболочки. Вирус является этиологическим агентом такого заболевания человека, как паралитический полиомиелит. Вирус инфицирует эпителий кишечника человека и в редких случаях центральную нервную систему, вызывая параличи (Eggers, 1999; Nathanson, 2008; Mach *et al.*, 2014).

Геном вируса полиомиелита представляет собой одноцепочечную РНК положительной полярности длиной примерно 7450 нуклеотидов. Вирусный геном кодирует полипротеин массой 247 кДа, который посттрансляционно расщепляется на функциональные вирусные белки (Pallansch et al., 1984). С 5'-концом генома ковалентно связан вирусный белок VPg (Lee et al., 1977), 3'-конец генома полиаденилирован (Yogo & Wimmer, 1972; Dorsch-Häsler, Yogo & Wimmer, 1975). Нетранслируемые области на 5' и 3' концах генома содержат структурированные участки РНК, регулирующие процессы репликации и трансляции генома, а также обуславливающие стабильность вирусной РНК в клетке: oriL (origin left), oriR (origin right), IRES (internal ribosome entry site) (Trono et al., 1988; Pelletier & Sonenberg, 1989; Pilipenko et al., 1996) (Рисунок 1). Регуляторные структурные элементы находятся и в кодирующей области генома. Достаточно хорошо изучены oril (origin internal, также называемый *cre*) в области генома, кодирующей белок 2С, и ингибитор РНКазы L в области генома, кодирующей белок 3С (Rieder et al., 2000; Goodfellow, Kerrigan & Evans, 2003b; Han et al., 2007). Не так давно открыты структурные элементы a и b в области генома, кодирующей белок 3D (Song et al., 2012; Burrill et al., 2013). Структурированные регуляторные элементы вирусного генома выполняют свои функции путем взаимодействия с различными вирусными и клеточными белками. Ключевая роль РНК-белковых взаимодействий в цикле репродукции делают вирус полиомиелита удобной моделью для изучения РНК-белкового узнавания. Этому также способствуют отработанные методы культивации и характеризации вируса, и его безопасность.



Рисунок 1. Схематичное представление генома полиовируса. Обозначены репликативные элементы *oriL, oriI, oriR*; размеры областей генома, кодирующих белки. Представлена вторичная структура репликативного элемента *oriL* (Ohlenschläger *et al.*, 2004)). Указан аминокислотная последовательность, вовлеченная во взаимодействие со стороны белка 3CD

## 1.1 Цикл репродукции вируса полиомиелита

Весь цикл репродукции полиовируса в культуре эукариотических клеток занимает около 7 часов и включает следующие этапы: 1) проникновение геномной РНК в клетку опосредованное взаимодействием с полиовирусным рецептором (клеточный трансмембранный гликопротеин CD155) (Mendelsohn, Wimmer & Racaniello, 1989; He *et al.*, 2000); 2) трансляция и репликация вирусного генома; 3) созревание вирусных частиц; 4) выход из клетки (Рисунок 2). В культуре клеток HeLa проникшая вирусная РНК детектируется в цитоплазме через 30-40 мин после адсорбции вируса, и через 1-1.5 часа после начала инфекции методом иммунофлуоресценции можно обнаружить вирусные белки и геномную РНК в цитоплазме в небольших комплексах (Egger & Bienz, 2002, 2005). Через 2-3 часа после начала инфекции, геномная РНК и вирусные белки (2B и, возможно, другие) локализуются на мембранах эндоплазматического ретикулума (Bienz, Egger & Pasamontes, 1987; Bolten *et al.*, 1998; Egger *et al.*, 2000; Egger & Bienz, 2002, 2005). Пик трансляции (3 часа после начала инфекции) совпадает по времени с появлением первых специфических для вирусной инфекции мембранных структур, на которых формируются множественные сайты репликации (Bienz *et al.*, 1987; Egger & Bienz, 2002; Belov *et al.*, 2012). Формирование мембранных структур позволяет защитить репликативные комплексы от внутриклеточных протеаз и нуклеаз, а также «спрятать» маркер вирусной инфекции – двуцепочечную РНК – от систем клеточного иммунитета. На ранних этапах инфекции эти структуры одномембранные, затем становятся двумембранными, усложняют свою форму, перемещаются в клетке в перинуклеарное пространство и образуют скопления - розетки (Bienz *et al.*, 1990; Troxler *et al.*, 1992; Egger *et al.*, 2000; Belov & Ehrenfeld, 2007; Belov *et al.*, 2007). Вместе с ними мигрирует геномная РНК и детектируемые вирусные белки (Bienz *et al.*, 1987; Bolten *et al.*, 1998; Egger *et al.*, 2000; Egger & Bienz, 2002, 2005).



Рисунок 2. Цикл репродукции вируса полиомиелита.

Считается, что за формирование «розеточных» структур ответственны вирусные белки 2ВС

и 2С, так как независимая экспрессия этих белков в эукариотических клетках приводит к формированию мембранных структур, морфологически схожих со структурами, образующимися при вирусной инфекции (Bienz *et al.*, 1990; Cho *et al.*, 1994; Aldabe & Carrasco, 1995; Barco & Carrasco, 1995; Aldabe, Barco & Carrasco, 1996; Teterina *et al.*, 1997).

Плюс цепей РНК синтезируется в среднем в 100 раз больше, чем минус цепей (Andino *et al.*, 1990b; Troxler *et al.*, 1992). Максимум синтеза (+) и (-) РНК наступает через 3-3,5 часа после начала инфекции (Troxler *et al.*, 1992; Bolten *et al.*, 1998; Belov *et al.*, 2012).

## 1.1.1 Проникновение вирусной РНК в клетку

Рецептор вируса полиомиелита CD155 представляет собой трансмембранный иммуноглобулин-подобный гликопротеид, вовлеченный в межклеточные взаимодействия (Mendelsohn *et al.*, 1989; Racaniello, 1996; Zhang *et al.*, 2008). Связывание с рецептором инициирует конформационные перестройки вирусной частицы, приводящие к выходу вирусной РНК из капсида и проникновению ее в клетку (He *et al.*, 2000, 2003; Zhang *et al.*, 2008). После проникновения вирусной РНК в клетку клеточный фермент 5'тирозил-ДНК-фосфодиэстераза 2 (TDP2) отщепляет белок VPg с 5'конца генома полиовируса (Langereis *et al.*, 2014).

## 1.1.2 Трансляция вирусной РНК и процессинг полипротеина

Трансляция вирусной РНК осуществляется в цитоплазме инфицированной клетки с помощью трансляционного аппарата хозяйской клетки. Однако процесс инициации трансляции у полиовируса от такового в эукариотических клетках. Геномные вирусные РНК не имеют структуры «кэп» (7mG) на 5'-конце, которая могла бы служить, как у клеточных мРНК, «якорем» для сборки трансляционного комплекса (Flanegan *et al.*, 1977; Lee *et al.*, 1977). Кроме того, вирусная протеаза 2A расщепляет эукариотический фактор инициации трансляции eIF4G и, таким образом, трансляция по каноническому пути в инфицированных клетках значительно подавлена (Gradi *et al.*, 1998; Novoa & Carrasco, 1999; Castello, Alvarez & Carrasco, 2006).

Вместо структуры кэп за инициацию трансляции вирусного генома отвечает сложно структурированный PHK-элемент IRES (internal ribosome entry site) (Pelletier *et al.*, 1988; Pelletier & Sonenberg, 1989). В инициацию трансляции вирусного генома помимо канонических факторов трансляции вовлечены различные клеточные белки, называемые ITAFs (IRES translation activation

factors): PCBP2 (также называемый hnRNP E – heterogeneous nuclear ribonucleoprotein E), PTB (polypyrimidine tract binding protein), PABP, SRp20 и другие белки (Hellen *et al.*, 1993; Walter *et al.*, 2002; Song *et al.*, 2005; Bedard, Daijogo & Semler, 2007). Показано, что клеточный белок PCBP2 связывается с IV доменом IRES и что это взаимодействие необходимо для инициации трансляции (Blyn *et al.*, 1996, 1997; Gamarnik & Andino, 1997; Walter *et al.*, 2002; Sean, Nguyan & Semler, 2008). Продемонстрировано положительное влияние комплекса PCBP/*oriL*, поли-А хвоста на 3'-конце генома и вирусного белка 2A на трансляцию вирусного генома в бесклеточной системе, в частности, на сборку и стабилизацию рибосом на вирусной PHK (Ogram *et al.*, 2010).

Синтезированный полипротеин расщепляется на вирус-специфические белки (Рисунок 3) (Pallansch et al., 1984). Кодирующую область вирусного генома можно условно разбить на три участка – P1, P2 и P3. Участок P1 кодирует белки капсида - VP1, VP2, VP3 и VP4. Участок P2 кодирует белки, вовлеченные в модификацию клеточной среды, - протеазу 2А, белок 2В, взаимодействующий с мембранами, и белок 2С, для которого предсказана функция хеликазы. Участок РЗ кодирует белки, вовлеченные в формирование репликативного комплекса, – белок ЗА, заякоривающий репликативный комплекс на мембранах; белок-затравку для синтеза дочерних цепей 3В; протеазу и РНК-связывающий белок 3С и РНК-зависимую РНК-полимеразу 3D. На начальных этапах инфекции перечисленные белки входят в состав полипротеинов, которые вовлекаются в формирование репликативных комплексов. Расщепление промежуточных полипротеинов, похоже, осуществляется на месте реализации функции включенных в них белков. Отщепление P1 происходит котрансляционно и осуществляется вирусной протеазой 2A (Toyoda et al., 1986). Отщепление полипротеина Р2 от полипротеина Р3 и их расщепление на функциональные белки осуществляется протеазой 3CD (Hanecak et al., 1982; Jore et al., 1988; Ypma-Wong et al., 1988) (Рисунок 3). Протеза 3CD также осуществляет расщепление полипротеина P1 на капсидные белки VP0, VP1 и VP3 (Ypma-Wong & Semler, 1987; Ypma-Wong et al., 1988).. Расщепление VP0 на VP2 и VP4 осуществляется автокаталитический в ходе созревания капсида (Harber et al., 1991; Basavappa et al., 1994). Протеаза 2А осуществляет альтернативный протеолиз 3CD, однако какой-либо значимой функции продуктов альтернативного протеолиза обнаружено не было (Toyoda *et al.*, 1986; Ventoso & Carrasco, 1995).



Рисунок 3. Схема процессинга полипротеина, (Parsley, Cornell & Semler, 1999) с модификациями. В скобках указаны молекулярные массы белков. Треугольниками отмечены сайты расщепления протеазой 3CD, ромбами отмечены сайты расщепления протеазой 2А. Звездой отмечен сайт автопроцессинга.

### 1.1.3 Сопряжение трансляции и репликации

Существуют данные, указывающие на невозможность инициации репликации на молекуле, не транслированной РНК. Помимо того, что вирусные белки непосредственно участвуют в репликации, показано, что некоторые белки функционируют именно на той молекуле РНК, которая служила матрицей для их трансляции (Bernstein, Sarnow & Baltimore, 1986; Johnson & Sarnow, 1991; Novak & Kirkegaard, 1994).

Во время трансляции геномная РНК не может быть использована в качестве матрицы для синтеза (-) цепи (Barton, Morasco & Flanegan, 1999). Эти два процесса идут в разных направлениях, и теоретически полимераза будет встречаться на матрице с рибосомами. Известно, что прекращения трансляции недостаточно для того, чтобы начался синтез (-) РНК, так как ингибирование трансляции пуромицином не стимулирует репликацию (Herold & Andino, 2001). Значит, необходимы дополнительные события для инициации репликации (Herold & Andino, 2001). Раз некоторые вирусные по-видимому, работают *in-cis* значит на каждой молекуле отдельно происходят события инициации одних процессов и терминации других. Соответственно, на каждой молекуле отдельно должно произойти событие прекращения инициации трансляции и событие инициации репликации (Barton *et al.*, 1999).

Согласно одной из принятых моделей, прекращение инициации трансляции обусловлено расщеплением белков ITAF (IRES trans acting factor) вирусными протеазами (Daijogo & Semler, 2011). Из этих белков наиболее изученным является белок PCBP2. Он имеет 3 домена, для которых показана PHK-связывающая активность, называемых KH-доменами (K-homology, см. список сокращений) (Siomi *et al.*, 1993, 1994; Perera *et al.*, 2007). Удаление третьего KH-домена лишает белок PCBP2 способности связывать IRES и другие белки (Srp20, PABP) и, соответственно, способности иницировать трансляцию (Perera *et al.*, 2007; Chase, Daijogo & Semler, 2014) (Pucyhok 4). Полное расщепление клеточного белка PCBP2 вирусной протеазой 3CD в инфицированной культуре клеток происходит через 3-4 часа после начала инфекции (Gamarnik & Andino, 1998; Perera *et al.*, 2007; Chase *et al.*, 2014). Именно в этот момент наблюдается пик синтеза дочерних PHK (Gamarnik & Andino, 1998; Perera *et al.*, 2007; Belov *et al.*, 2012; Chase *et al.*, 2014). Подавление расщепления PCBP2 ингибирует вирусную репликацию (Chase *et al.*, 2014). Таким образом, расщепление клеточного белка инициации репликации. Остается, однако, непонятным, каким образом дочерние копии PHK будут вовлекаться в трансляцию в отсутствие целого PCBP2.



Рисунок 4. Модель переключения процессов трансляции и репликации через расщепление PCBP2 (по Chase et al., 2014, с модификациями).

Другие ITAF также расщепляются вирусными протеазами, но позднее. Например, через 7 часов после инфекции в культуре клеток расщепляется более 60% белка PABP (Joachims, Van Breugel & Lloyd, 1999; Kuyumcu-Martinez *et al.*, 2004; Bonderoff, Larey & Lloyd, 2008). Однако какое

влияние может оказывать это событие на процессы вирусной репродукции на поздних сроках инфекции неясно.

## 1.1.4 Репликация вирусного генома

Непосредственно синтез дочерних копий РНК производит вирусная РНК-зависимая РНКполимераза 3D (Baltimore et al., 1963). Координация этапов репликации осуществляется через взаимодействие вирусных и клеточных белков с репликативными РНК-элементами вирусного генома – oriL, oriI (cre) и oriR. На данный момент в литературе обсуждаются три основных комплекса, рибонуклеопротеидных образующихся в ходе репликации. Это комплекс oriL/3CD/3AB/PCBP2 на 5'конце генома, который, который вовлечен в закрепление репликативного комплекса на мембранных структурах, инициацию синтеза (-) и (+) цепей (Xiang et al., 1995а; Gamarnik & Andino, 1997, 2000; Silvera, Gamarnik & Andino, 1999). Второй комплекс - oriI/3CD в области, кодирующей белок 2С, отвечающий за уридилирование вирусного белка VPg (3B), который в последствии будет служить затравкой для РНК зависимой РНК полимеразы предположительно в ходе синтеза (+) цепей и, возможно, (-) цепей (Rieder et al., 2000; Goodfellow et al., 2003b; Morasco et al., 2003; Murray & Barton, 2003; Yin et al., 2003; Shen et al., 2008). Третий комплекс oriR/3CD/polyA/PABP формируется на 3'конце генома и участвует в его циркуляризации через взаимодействие с комплексом на 5'-конце (Herold & Andino, 2001). VPg, инициация (-) PHK. Здесь описание известных РНК-белковых комплексов и их возможных ролей в ходе репликации. Обсуждение того, что нет окончательного понимания, где уридилируется VPg для того, чтобы стать затравкой для минус цепи будет позже.

Формирование репликативного комплекса необходимо для пространственной организации трех элементов, необходимых для синтеза новых цепей: полимеразы, матрицы для синтеза и затравки. Еще одним необходимым элементом для репликации является белок 2С – АТФаза, PHKсвязывающий и мембран-связывающий белок – самый известный ингибитор репликации полиовируса гуанидин гидрохлорид ингибирует функцию именно этого белка (Tolskaya *et al.*, 1994; Barton & Flanegan, 1997; Pfister & Wimmer, 1999). Существует следующая гипотетическая модель сборки и функционирования репликативного комплекса на мембранах: полипротеины ЗАВ (содержит белок-затравку для полимеразы 3В или VPg) и 2BC (содержит белок 2С), связываются мембранами и привлекают к ним геномную PHK (матрицу) и 3CD (полипротеин-предшественник полимеразы 3D) напрямую или в составе длинного предшественника и через взаимодействие с репликативными элементами генома инициируют репликацию (Collis *et al.*, 1992; Rodríguez &

19

Carrasco, 1993, 1995; Harris et al., 1994; Lama, Sanz & Rodríguez, 1995; Teterina et al., 1997; Sharma et al., 2009). В пользу этой гипотезы говорит ряд фактов: белок ЗАВ взаимодействует с полимеразой 3D и стимулирует ее активность в 75-100 раз (Lama et al., 1994; Plotch & Palant, 1995; Hope, Diamond & Kirkegaard, 1997; Strauss & Wuttke, 2007; DeStefano, 2010), белок ЗАВ взаимодействует с полипротеином 3CD в составе комплексов на репликативных элементах oriL и oriR и стимулирует расщепление 3CD на 3C и 3D (Harris et al., 1994; Molla et al., 1994; Xiang et al., 1995a, 1995b, 1998), протеаза 3CD отщепляет затравку 3В только от белка 3AB, ассоциированного с мембранами (Lama et al., 1994), и белок ЗАВ связывается с 2С (Teterina et al., 2011). Существуют экспериментальные данные, указывающие на то, что белки ЗАВ и 2ВС не могут функционировать in trans (т.е. если экспрессированы на другой матрице) (Bernstein et al., 1986; Johnson & Sarnow, 1991; Collis et al., 1992; Novak & Kirkegaard, 1994), хотя в случае ЗАВ функция in trans может осуществляться более длинным предшественником (Giachetti, Hwang & Semler, 1992; Towner, Mazanet & Semler, 1998; Liu et al., 2007). Этот эффект может иметь несколько объяснений, в том числе: 1) сборка репликативного комплекса на матричной РНК инициируется синтезированными на этой РНК белками, 2) после формирования репликативного комплекса, другие белки не могут в него проникнуть. Несмотря на продолжительную историю изучения репликации вируса полиомиелита, детали формирования репликационных комплексов и сам процесс репликации остается мало изученным до сих пор. Тем не менее, накопились данные, которые позволяют сформировать более-менее цельную картину. Рассмотрим их подробнее.

#### 1.1.4.1 Рибонуклеопротеидный комплекс на 5'-конце генома

Репликативный элемент *ori*L на 5'конце полиовирусного генома называют клеверным листом из-за характерной вторичной структуры – элемент разделяют на 4 домена: ножку - домен *a* и три шпильки - домены *b*, *c* и *d* (Рисунок 1). Известно, что с репликативным элементом oriL взаимодействуют вирусные белки 3CD и 3AB (Andino *et al.*, 1990a; Harris *et al.*, 1994; Xiang *et al.*, 1995b; Blair *et al.*, 1996) и клеточные белки PCBP2 и p36 (фрагмент фактора элонгации EF1a) (Harris *et al.*, 1994; Parsley *et al.*, 1997, 1999; Gamarnik & Andino, 2000). Мутации в oriL, нарушающие взаимодействие с PCBP2 и 3CD, приводят к быстрому разрушению вирусной PHK, по крайней мере, в бесклеточной системе (Barton, O'Donnell & Flanegan, 2001; Kempf & Barton, 2008a).

Значение домена d для инициации синтеза (-) цепи отмечается в различных работах (Andino *et al.*, 1990a; Barton *et al.*, 2001; Du *et al.*, 2003). Этот участок oriL узнается вирусными белками 3CD/3C у различных энтеровирусов (Andino *et al.*, 1990a; Leong, Walker & Porter, 1993; Zell *et al.*,

2002; Ohlenschläger *et al.*, 2004; Shih, Chen & Wu, 2004). Впервые, указание на взаимодействие 3CD и oriL было получено в результате возникновения компенсаторных аминокислотных замен Thr<sub>154</sub>lle и Lys<sub>156</sub>Arg в 3C при вставке 4 нуклеотидов после 67-го нуклеотида вирусного генома в домене *d* (Andino *et al.*, 1990b). Вирусный белок 3CD был обнаружен в комплексах с oriL в лизатах инфицированной культуры HeLa с помощью антител, специфичных к 3C или к 3D (Andino *et al.*, 1990a). Хотя белок 3CD полиовируса эффективнее чем 3C связывается с oriL, конкретные аминокислотные последовательности, вовлеченные во взаимодействие, известны только для 3C (Harris *et al.*, 1994). Белок 3CD является предшественником вирусной PHK-зависимой PHK-полимеразы 3D, которая осуществляет синтез (-) и (+) PHK и уридилирование VPg (Baltimore *et al.*, 1963; Baltimore, 1964; Paul *et al.*, 1998; Goodfellow *et al.*, 2003b; Morasco *et al.*, 2003; Shen *et al.*, 2008). Также белок 3CD имеет протеолитическую активность: участвует, как уже отмечалось выше, в процессинге капсидных белков, полипротеинов P2 и P3, а также расщепляет ряд клеточных белков (в частности PCBP) (Rochl *et al.*, 1997; Perera *et al.*, 2007).

Домен b oriL оказывает влияние на синтез (+) и (-) цепей. С доменом b репликативного элемента oriL и полипиримидиновым трактом, расположенным сразу после oriL (нуклеотиды 92-102), связываются клеточные белки PCBP 1 и 2 (Roehl *et al.*, 1997; Gamarnik & Andino, 1998; Walter *et al.*, 2002; Perera *et al.*, 2007). Способность PCBP связываться с oriL и стимулировать репликацию вирусной PHK в бесклеточной системе сохраняется после отщепления протеазой 3CD домена KH3, необходимого для связывания белка PCBP2 с IRES (Walter *et al.*, 2002; Perera *et al.*, 2007). Многие авторы считают, что формирование комплекса oriL/3CD/PCBP необходимо для инициации репликации (Harris *et al.*, 1994; Gamarnik & Andino, 1997, 1998; Parsley *et al.*, 1997; Silvera *et al.*, 1999; Herold & Andino, 2001; Walter *et al.*, 2002). В других работах показано, что если защитить 5'-конец вирусного генома с помощью кэп-структуры, то отсутствие PCBP скажется только на эффективности трансляции, но не на эффективности репликации в бесклеточной системе (Murray, Roberts & Barton, 2001; Murray & Barton, 2003; Kempf & Barton, 2008a, 2008b). Таким образом, положительное влияние PCBP на репликацию в экспериментальных условиях может быть обусловлено тем, что образующийся на 5'-конце геномной PHK комплекс защищает PHK от атаки экзонуклеаз (Murray *et al.*, 2001; Murray & Barton, 2003; Kempf & Barton, 2008a, 2008b).

Рассмотрим роль домена *а* элемента oriL. В работе Vogt и Andino (2010) использовалась полногеномная РНК вируса полиомиелита, несущая два элемента *oriL*. Один служил для инициации синтеза (-) цепи, но не мог поддерживать синтез (+) цепи. Домен *а* второго элемента модифицировался для исследования его роли в синтезе (+) цепи. В результате было показано, что

последовательность домена *a* (а именно вторая и третья A-U пары спирали) важна для инициации синтеза (+) цепи и что этот участок не влияет на инициацию синтеза (-) цепи.

#### 1.1.4.2 Уридилирование VPg

Затравкой для синтеза дочерних цепей служит диуридилированный (несущий два нуклеотида U, VPgpUpU<sub>OH</sub>) белок 3B (или VPg) (Paul *et al.*, 1998). Таким образом, дочерние (-) и (+) цепи имеют белок VPg на 5'-конце (Flanegan *et al.*, 1977; Lee *et al.*, 1977). Белок 3B рекрутируется в репликативный комплекс скорее всего в виде предшественника - 3AB или в виде более длинного предшественника полипротеина P3 (Liu *et al.*, 2007; Cameron, Oh & Moustafa, 2010; Spear *et al.*, 2015). Последнее подтверждается тем, что мутации в гидрофобном домене белка 3A (в составе 3AB), приводящие к дефектам репликации, не могут быть компенсированы белком 3AB без мутаций, но могут быть компенсированы более длинным предшественником 3AB - полипротеином 3ABCD (P3) (Giachetti *et al.*, 1992; Towner *et al.*, 1998; Liu *et al.*, 2007).

Уридилирование VPg может осуществляться только на матрице (Takeda *et al.*, 1987). В принципе, уридилирование VPg может осуществляться на любой поли-А последовательности (Paul *et al.*, 1998). В геноме полиовируса есть по крайней мере две предполагаемые последовательности для уридилирования VPg. Это поли-А на 3'конце генома и элемент отіІ в области, кодирующей 2С. До сих пор нет единого мнения, какая из этих последовательностей служит матрицей для уридилирования VPg в синтезе (-) цепей. Изначально предполагаали, что уридилирование на отіІ необходимо для синтеза и (-) и (+) цепей (Goodfellow *et al.*, 2000; Paul *et al.*, 2000). В более поздних работах показано, что геномы без отіІ способны к инициации синтеза (-) цепи, но не могут синтезировать дочерние (+) цепи (Goodfellow *et al.*, 2003а; Morasco *et al.*, 2003; Murray & Barton, 2003). Было высказано предположение, что синтез (-) цепи начинается уридилированием VPg на поли-А последовательности на 3'-конце генома (Goodfellow *et al.*, 2003а; Morasco *et al.*, 2003; Murray & Barton, 2003; Silvestri *et al.*, 2006).

Уридилирование на матрице oril в бесклеточной системе стимулируется в присутствии белка 3CD (Paul *et al.*, 2000; Goodfellow *et al.*, 2003b). В работе Shen *et al.* (2008) было показано, что на oril формируется комплекс из двух молекул 3C и одной молекулы 3D, осуществляющий уридилирование VPg. Была предложена трехмерная модель этого комплекса и высказано предположение, что димер 3C разрушает спираль под функциональной петлей oril (Amero *et al.*, 2008; Shen *et al.*, 2008). Показано также, что для уридилирования VPg для синтеза (+) цепей

оказывается важным формирование комплекса 3CD/oriL (Vogt & Andino, 2010).

Таким образом, oril необходим для инициации синтеза дочерних (+) цепей, но пока не совсем очевидна его роль в синтезе (-) цепи. Показано, что положение oril в геноме не имеет решающего значения, и мутации в исходном oril могут быть компенсированы oril-подобным элементом в другом участке генома, что свидетельствует в пользу того, что этот репликативный элемент используется *in* cis (Yin *et al.*, 2003).

Существует альтернативная гипотеза, предполагающая уридилирование VPg для синтеза (+) цепи на 3' конце (-) цепи, а именно на последовательности 3'AAUUUUGUC (Brunner *et al.*, 2005; Sharma *et al.*, 2005). В поддержку этой гипотезы выступают данные о формировании рибонуклеопротеидного комплекса на *oriL* на 3'-конце (-) цепи (Brunner *et al.*, 2005). Однако, предположение о том, что oriL на 3'конце (-) цепи имеет какую-либо функцию вступают в противоречие с данными других работ, где разрушение вторичной структуры на (-) цепи при сохранении вторичной структуры на (+) цепи не оказывает влияние на эффективность синтеза (+) цепей (Vogt & Andino, 2010).

### 1.1.4.3 Рибонуклеопротеидный комплекс на 3'-конце генома

На 3' конце полиовирусного генома расположена нетранслируемая область, состоящая из 68 нуклеотидов и представляющая собой репликативный элемент oriR. На 3'конце после oriR распологается последовательность поли-А. У полиовируса первого типа штамма Mahoney, вируса коксаки ВЗ и некоторых других вирусов длина поли-А последовательности составляет 80-90 остатков (Spector & Baltimore, 1974).

Репликативный элемент oriR генома полиовируса состоит из двух шпилек (X и Y). Петли шпилек комплементарно взаимодействуют друг с другом, формируя так называемый «киссинг» (Pilipenko *et al.*, 1996) (Рисунок 5). Поли-А может быть вовлечен в формирование дополнительного домена oriR (домен S) (Pilipenko *et al.*, 1992b, 1996)



Рисунок 5. Третичная структура oriR вируса полиомиелита. Линиями обозначены взаимодействия нуклеотидов петель X и Y. Цифрами указаны номера нуклеотидов генома начиная от 5' конца. (Pilipenko et al., 1996)

Нарушение вторичной структуры oriR у полиовируса ингибирует репликацию (Pilipenko *et al.*, 1996; Brown *et al.*, 2005). При этом удаление oriR не приводит к утрате инфекционности, не влияет на эффективность трансляции или стабильность полногеномных транскриптов, хотя и снижает их инфекционность на 6 порядков (после 3 пассажей после трансфекции вирус восстанавливает фенотип и приобретает аминокислотные замены в 2C и 3D) (Todd *et al.*, 1997; Brown *et al.*, 2005). Интересно отметить, что замена oriR у полиовируса на таковой вируса коксаки B4 или риновируса 14 практически не влияет на эффективность репродукции, хотя структуры oriR у этих вирусов различаются (Rohll *et al.*, 1995).

Длина поли-А хвоста геномной РНК влияет на эффективность синтеза (-) цепи. Предполагалось, что этот эффект обусловлен необходимостью взаимодействия с РАВР (Herold & Andino, 2001; Silvestri *et al.*, 2006). Однако полногеномная вирусная РНК полиовируса эффективно реплицируется в бесклеточной системе обедненной методом иммунопреципитации по белку РАВР, хотя геномы без поли-А на это не способны (Svitkin *et al.*, 2007). Значит, наличие РАВР не играет решающей роли и функция поли-А в репликации связана не столько со взаимодействием с РАВР, но, возможно, с пространственной организации циркуляризации генома (Svitkin *et al.*, 2007).

У вирусного генома поли-А хвост обычно составляет 80-90 нуклеотидных остатков, тогда как у дочерней (-) цепи поли-U составляет около 20 остатков (Steil, Kempf & Barton, 2010). Для другого представителя рода энтеровирусов вируса коксаки ВЗ показано, что синтез поли-А хвоста правильной длины возможен только при наличии функционального репликативного элемента oriR на 3'-конце вирусного генома (van Ooij *et al.*, 2006). Чтобы объяснить, каким образом синтезируется длинная поли-А последовательность на дочерних (+) цепях, если матрица поли-U длиной всего около 20 нуклеотидных остатков, предложена модель повторов транскрипции полимеразой 3D, в которой при синтезе поли-А полимераза останавливается, дуплекс расплетается, комплекс элонгации сдвигается к 3'-концу матрицы и синтез начинается снова (Steil *et al.*, 2010; Kempf *et al.*, 2013).

Таким образом, элементы на 3'конце не видоспецифичны и не могут надежно контролировать выбор матрицы для начала репликации. В свете того, что на 3'конце генома должен осуществляться старт синтеза (-) цепи, такой вывод представляется парадоксальным (Agol, Paul & Wimmer, 1999). По-видимому, специфическим регулятором является репликативный элемент на 5'-конце вирусного генома. Тогда каким образом сформированный репликативный комплекс с 5'конца генома попадет на 3'конец для начала синтеза (-) цепи?

Белок 3CD взаимодействует и с oriR, и с oriL (Harris, Xiang, et al., 1994). Также показано, что 3CD связывается с PCBP и PABP, и что PCBP и PABP взаимодействуют друг с другом (Herold & Andino, 2001). Это позволило предположить, что комплексы PABP/поли-A и 3CD/oriR на 3'конце взаимодействуют с комплексом oriL/3CD/PCBP на 5'конце генома и, таким образом, обеспечивают его циркуляризацию (Barton *et al.*, 2001; Herold & Andino, 2001), как это происходит у клеточных мPHK (Ostareck-Lederer, Ostareck & Hentze, 1998; Wang *et al.*, 1999) (Рисунок 6). Нарушение этих взаимодействий негативно сказывается на инициации репликации (Herold & Andino, 2001).



Рисунок 6. Схема циркуляризации генома у клеточных мРНК (А) и полиовирусного генома (В) (Herold et al., 2001).

## 1.1.5 Сборка капсида и упаковка дочерних молекул (+)РНК

Полиовирус имеет икосаэдрический капсид диаметром около 30 нм и не имеет липидной оболочки. Снаружи капсид образован белками VP1, VP2, и VP3 (массой 32, 29 и 26 кДа). Изнутри он выстлан белком VP4 (7 кДа) и N-концевыми участками белков VP1, VP2 и VP3. Каждый белок представлен 60 копиями. Сборка капсида осуществляется поэтапно: сначала из белков VP1, VP0 и VP3 собирается протомер, затем пять протомеров формируют пентамер, 12 пентамеров формируют целый капсид (Hogle, Chow & Filman, 1985) (Рисунок 7). В ходе созревания капсида VP0 автокаталитически расщепляется на VP4 и VP2.



Рисунок 7. Структура капсида полиовируса (по Hogle et al., 1985).

Процесс упаковки дочерних молекул РНК и формирования новых вирусных частиц изучен слабо. Показано, что процессы репликации и инкапсидации сцеплены друг с другом, вероятно, для обеспечения специфичности упаковки (Nugent *et al.*, 1999).

В процесс упаковки геномных РНК вовлечены вирусные белки. В настоящий момент установлено, что вирусный белок 2С участвует в процесс морфогенеза вирусных частиц (Li & Baltimore, 1988, 1990; Vance *et al.*, 1997; Wang *et al.*, 2012, 2014; Asare *et al.*, 2016). Наиболее прямым свидетельством того, что 2С участвует в инкапсидации, служит непосредственное взаимодействие 2С и капсидного белка VP3 (Liu *et al.*, 2010). Кроме того, есть указания, что вирусные белки 3CD и 3AB, возможно, участвуют в упаковке дочерних геномов (Oh *et al.*, 2009).

## 1.2 Взаимное узнавание 3CD и *ori*L: структурно-функциональная характеристика

## 1.2.1 Белок 3CD

Вирусный белок 3CD является предшественником цистеиновой протеазы 3C (Ivanoff *et al.*, 1986; Mosimann *et al.*, 1997) и PHK-зависимой PHK-полимеразы 3D (Flanegan & Van Dyke, 1979; Neufeld, Richards & Ehrenfeld, 1991; Gohara *et al.*, 2000). Белок 3CD обладает протеолитической и PHK-связывающей активностью, но не обладает полимеразной активностью (Harris *et al.*, 1992;

Тhompson & Peersen, 2004). В свободной полимеразе 3D N-концевой участок завернут внутрь молекулы 3D и связан там водородными связями (Hansen, Long & Schultz, 1997). Такое расположение N-концевого участка необходимо для нормального функционирования полимеразы 3D (Thompson & Peersen, 2004). В полипротеине 3CD N-конец 3D связан с C-концом 3C и, таким образом, не способен заворачиваться в молекулярный карман 3D (Marcotte *et al.*, 2007). Этим объясняется отсутствие полимеразной активности у 3CD, несмотря на наличие полимеразного домена (Thompson & Peersen, 2004). Таким образом, отщепление 3D необходимо для инициации репликации. Белок 3CD способен расщеплять себя *in trans* и, возможно, *in cis* на 3C и 3D (Hanecak *et al.*, 1984; Ivanoff *et al.*, 1986). Интересно отметить, что существуют замены в белке 3C (валин в положении 54 на аланин или изолейцин в положении 74 на треонин), которые с одной стороны приводят к потере протеазой способности расщеплять полипротеин 3CD на 3C и 3D и эффективно реплицироваться, а с другой стороны не сказываются на эффективности расщепления P1 на белки капсида (Dewalt & Semler, 1987; Kean *et al.*, 1988). Таким образом, протеолитическая активность 3CD регулируется по-разному в зависимости от субстрата.

В кристалле 3CD структура доменов 3C и 3D практически никак не отличаются от структур 3C и 3D, закристаллизованных отдельно, за исключением C-конца 3C и N-конца 3D, которые у свободных молекул приобретают другую конформацию (Hansen *et al.*, 1997; Mosimann *et al.*, 1997; Marcotte *et al.*, 2007) (Рисунок 8). Однако протеолитическая и PHK-связывающая активность 3CD отличается от таковых для 3C (Burns *et al.*, 1989; Parsley *et al.*, 1999). Замена второй трети аминокислотной последовательности полимеразы 3D полиовируса на последовательность полимеразы 3D вируса коксаки B3 нарушает способность 3CD образовывать комплекс с oriL (Parsley *et al.*, 1999; Cornell & Semler, 2002). Эти данные указывают на вовлечение домена 3D белка 3CD в узнавание oriL. Домен 3D белка 3CD может участвовать в узнавании oriL несколькими способами (либо их комбинацией): 1) взаимодействие домена D и домена C внутри одной молекулы 3CD, 2) межмолекулярное взаимодействие двух молекул 3CD друг с другом, 3) прямое узнавание oriL аминокислотами домена 3D белка 3CD.



Рисунок 8. Схематическое представление кристаллической структуры 3CD (код в базе данных PDB - 2IJD) (Marcotte *et al.*, 2007). Отмечены домен 3D, домен 3C и линкер, их соединяющий.

Взаимодействие между поверхностями 3D и 3C в молекуле 3CD продемонстрировано экспериментально (Moustafa *et al.*, 2015). Методами молекулярной динамики и малоуглового рентгеновского рассеивания показано, что в растворе полипротеин 3CD может находится примерно с одинаковой вероятностью в двух конформациях – растянутой, как в кристалле, где домен 3C и домен 3D связаны только линкером (Pucyhok 8; (Marcotte *et al.*, 2007)), и компактной, где образуется поверхность взаимодействия между доменами 3C и 3D (Moustafa *et al.*, 2015). В компактной форме нуклеотидный карман 3D закрыт, что, как считают авторы, является еще одним препятствием для работы полимеразного домена в составе 3CD (Moustafa *et al.*, 2015).

Необходимость РНК-связывающей активности 3С и полимеразной активности 3D для инициации репликации показана в экспериментах с выделенными репликативными комплексами (Spear *et al.*, 2015). А именно: с oriL должны взаимодействовать две молекулы белка 3CD (или предшественника P3), одна из которых должна обладать PHK-связывающей активностью, а вторая содержать функциональную полимеразу (Spear *et al.*, 2015). Возможно, в комплексе с oriL, молекулы 3CD также взаимодействуют друг с другом, и связывание с oriL стимулирует отщепление активной полимеразы от 3CD. В поддержку этой гипотезы говорит несколько фактов. Экспрессируемая в эукариотических клетках активная протеаза 3CD не способна полностью расщепляться на 3C и 3D (Porter *et al.*, 1993). Компактность молекулы 3C подразумевает кооперативность структурных изменений в разных участках молекулы. Действительно, модель взаимодействия шпильки d *oriL* и 29

3C, полученная с помощью малоуглового рентгеновского рассеивания для родственного полиовирусу риновируса 2, демонстрирует структурные перестройки в протеолитическом сайте 3C при взаимодействии с PHK (Claridge *et al.*, 2009).

Известны аминокислотные последовательности и отдельные аминокислоты в 3С, замены в которых приводят к нарушению взаимного узнавания 3CD и *ori*L (Таблица 1). Они расположены в молекуле 3C на противоположенной стороне от протеолитического сайта (Mosimann *et al.*, 1997; Claridge *et al.*, 2009). Основными являются последовательности  $_{154}$ TGK $_{156}$  (Andino *et al.*, 1990а, 1990b) и  $_{82}$ KFRDI<sub>86</sub> (Hämmerle, Molla & Wimmer, 1992; Andino *et al.*, 1993) (Рисунок 9, .). Показано, что замены Thr $_{154}$ IIe и Lys $_{156}$ Arg в последовательности TGK могут компенсировать мутации в апикальном участке домена d *oriL* (Andino *et al.*, 1990b). Мотив  $_{82}$ KFRDIR $_{87}$  является абсолютно консервативным среди энтеровирусов. Этот мотив любопытен тем, что в нем чередуются аминокислоты, вовлеченные в протеолиз (Phe $_{83}$ , Asp $_{85}$ ), и аминокислоты, вовлеченные в связывание PHK (Lys $_{82}$ , Arg $_{84}$ , Asp $_{85}$ , Ile $_{86}$ ) (Hämmerle *et al.*, 1992; Andino *et al.*, 1993).

Структура 3С образована двумя β-бочонками, которые связаны друг с другом линкером (Mosimann *et al.*, 1997). Каждый бочонок образован двумя β-листами, образованными 6-ю β-тяжами - с а по f. Один β-тяж может участвовать в образовании обоих β-листов одного бочонка, например, e(1)I в одном β-листе переходит в e(2)I второго β-листа бочонка I, аналогично b1II переходит в b2II в домене II, тоже для bI и для eII. Также между бочонками на поверхности белка расположены 4 αспирали – A, B, C и D соответственно (Рисунок 9).

Аминокислотные остатки, участвующие в узнавании РНК, располагаются на петлях, связывающих некоторые  $\beta$ -тяжи, на спиралях С и D и на линкере (Рисунок 9). На петле, связывающей  $\beta$ -тяжи dII и e2II, располагается последовательность  $_{154}$ TGK $_{156}$  (Mosimann *et al.*, 1997). Мотив  $_{82}$ KFRDI $_{86}$  расположен частично в линкере (аминокислоты Lys $_{82}$ , Phe $_{83}$ , Arg $_{84}$ , Asp $_{85}$ , Ile $_{86}$ ), и частично в спирали С (Ile $_{86}$ , Arg $_{87}$ , Pro $_{88}$ , His $_{89}$ ).



Рисунок 9. Схематическое представление пространственной структуры протеазы 3С полиовируса. На рисунке обозначены аминокислоты, участвующие в связывании с РНК, и некоторые другие аминокислотные остатки: синим подчеркнут мотив KFRDI, зеленым последовательность TGK. Также обозначены α-спирали A, C и D; β-тяжи e11 (красным), e2I, b1II, b2II, e2II и dII. Указан линкер. (адаптировано из Melchers et al., 2006).

Аминокислоты белка ЗС	Известные функции аминокислот и эффекты аминокислотных замен в 3С*
Tyr <sub>6</sub> , Lys <sub>12</sub> , Arg <sub>13</sub>	взаимодействие с oril (Amero <i>et al.</i> , 2008) взаимодействие с <i>oriL</i> (Blair <i>et al.</i> , 1998)
His <sub>31</sub>	протеолиз Р1 и связывание с <i>oriL</i> (Andino <i>et al.</i> , 1993) поверхность внутримолекулярного взаимодействия в 3CD (Moustafa <i>et al.</i> , 2015)

Таблица 1 Функции аминокислот белка 3С

Asp <sub>32</sub>	взаимодействие с oriI (Amero <i>et al.</i> , 2008) взаимодействие с <i>oriL</i> (Andino <i>et al.</i> , 1993) поверхность внутримолекулярного взаимодействия d 3CD (Moustafa <i>et al.</i> , 2015)
His <sub>40</sub> , Glu <sub>71</sub> , Cys <sub>147</sub>	протеолитическая триада His40-Cys147-Glu71 (не влияют на взаимодействие с PHK) (Hämmerle <i>et al.</i> , 1991)
Lys <sub>52</sub>	аминокислота К <sub>52</sub> влияет на специфичность протеолиза (Dewalt & Semler, 1987)
Va <sub>154</sub> , Ile <sub>74</sub>	замена V54A приводит к тому, что 3CD не процессируется на 3C и 3D при сохранении эффективного протеолиза P1 (Dewalt & Semler, 1987) замена I74T снижает эффективность расщепления 3CD на 3C и на 3D и подавляет синтез дочерних PHK при сохранении эффективного протеолиза P1 (Kean <i>et al.</i> , 1988)
Asn <sub>80</sub> , Glu <sub>81</sub> , Arg <sub>87</sub>	взаимодействие с oril (Amero et al., 2008)
<sub>82</sub> KFRDI <sub>86</sub>	консервативный мотив у пикорнавирусов - аминокислоты K <sub>82</sub> , R <sub>84</sub> , I <sub>86</sub> важны для репликации (Hämmerle <i>et al.</i> , 1992) - аминокислоты F <sub>83</sub> , D <sub>85</sub> , R <sub>87</sub> важны для протеолитической активности (Hämmerle <i>et al.</i> , 1992) - аминокислоты R <sub>84</sub> , D <sub>85</sub> важны для взаимодействия с PHK (Andino <i>et al.</i> , 1993; Shen <i>et al.</i> , 2008)
His <sub>89</sub>	взаимодействие с oril (Amero <i>et al.</i> , 2008) слабое влияние на связывание с <i>oriL</i> (Blair <i>et al.</i> , 1998)
Tyr <sub>142</sub>	протеолитическая активность и взаимодействие с РНК (Blair et al., 1996)
154TGK156	аминокислоты TGK участвуют во взаимодействии с <i>oriL</i> (Andino <i>et al.</i> , 1993) замены Thr <sub>154</sub> Ile/Val и Lys <sub>156</sub> Arg компенсируют нарушения в апикальном участке домена d <i>oriL</i> (Andino <i>et al.</i> , 1990b) триплет <sub>154</sub> TGK <sub>156</sub> участвует в формировании интерфейса между 3C и 3D в димере в кристалле 3CD (Marcotte <i>et al.</i> , 2007) замена K <sub>156</sub> A подавляет способность 3CD стимулировать уридилирование VpG (Marcotte <i>et al.</i> , 2007) замены G <sub>155</sub> A и K <sub>156</sub> Q ослабляют протеолитическую активность 3C (EV71) (Shih <i>et al.</i> , 2004)
Gly <sub>149</sub> , Ala <sub>171</sub>	протеолитическая активность и взаимодействие с РНК (Andino et al., 1993)

His <sub>161</sub>	протеолитическая активность (Blair et al., 1996)
His <sub>161</sub> , Gly <sub>163</sub> , Gly <sub>164</sub>	узнавание белкового субстрата (Blair et al., 1996)
Val <sub>162</sub>	взаимодействие с oriI (Amero <i>et al.</i> , 2008)
Ala <sub>172</sub>	узнавание белкового субстрата (Blair <i>et al.</i> , 1996) взаимодействие с <i>oriL</i> (Blair <i>et al.</i> , 1996)
Arg <sub>176</sub>	связывание <i>oriL</i> (но не oriI) (Andino <i>et al.</i> , 1993; Blair <i>et al.</i> , 1998; Amero <i>et al.</i> , 2008) поверхность внутримолекулярного взаимодействия в 3CD (Moustafa <i>et al.</i> , 2015)
аминокислоты в положениях с 176 по185	внутримолекулярное взаимодействие в 3CD (Moustafa <i>et al.</i> , 2015)

\* - у полиовируса первого серотипа

## 1.2.2 Репликативный элемент oriL: характеристика домена d

Вторичная структура домена *d* oriL консервативна у рино- и энтеровирусов (Рисунок 10). Домен *d* клеверного листа представляет собой РНК-шпильку, образованную двумя участками спаренных нуклеотидов, представляющих собой А-спирали с выпетливанием между ними и апикальной шпилечной петлей, образованной тремя, четырьмя или пятью нуклеотидами (Рисунок 10).



Рисунок 10. Вторичная структура доменов d *oriL* у ряда представителей энтеровирусов (PV1 – полиовирус серотипа 1, CVB3 – коксакивирус B3, BEV1 – первый oriL бычьего энтеровируса 1, HRV2 и 14 – риновирус 2 и риновирус 14); жирным шрифтом отмечены нуклеотиды, предположительно вовлеченные во взаимодействие с 3C/3CD (Ohlenschläger *et al.*, 2004; Ihle *et al.*, 2005) (по Zell et. al, 2002).

Были исследованы комплексы доменов d с протеазой 3С коксакивирус B3 и бычьего энтеровируса 1 методом ЯМР (Ohlenschläger *et al.*, 2004; Ihle *et al.*, 2005). Сравнение данных ЯМР для свободного домена d и для комплекса с 3С позволили выявить нуклеотиды, потенциально взаимодействующие с 3С. Ими оказались нуклеотиды тетрапетли и ее фланкирующей пары, а также цитидины спирали в районе выпетливания (Рисунок 11).



Рисунок 11. Схематичное представление пространственной структуры домена d *oriL* коксакивируса ВЗ. Большой и малый желобки домена d и основание второго нуклеотидного остатка, выпетленное в малый желобок тетрапетли, отмечены стрелками. Основания нуклеотидных остатков, вовлеченные во взаимодействие с белком 3С, отмечены серым (по Ohlenschläger *et al.*, 2004).

Исследования с использованием рекомбинантного белка 3С коксакивируса B3 и *oriL* различных энтеровирусов продемонстрировали, что количество нуклеотидов в апикальной шпилечной петле домена d важнее для эффективности взаимодействия с 3С, чем ее последовательность (Zell *et al.*, 2002). Генетические эксперименты с химерами полиовируса 1 и риновируса 14 позволили предположить, что при связывании значение имеет именно пространственная структура апикального участка, т.е. ориентация структурных элементов, таких как рибозы и основания, друг относительно друга (Rieder *et al.*, 2003).

Методом ЯМР была выявлена пространственная структура апикального участок домена d для ряда энтеровирусов. Оказалось, что тетрапетли uCACGg (CVB3), uUACGg (HRV2), cUAUg (HRV14), cGUUAg (BEV1), uGCUAg (вариант PV1) имеют схожие структурные свойства, по которым их можно отнести к пространственному структурному классу тетрапетель UNCG (N= любой нуклеотид) (Du *et al.*, 2003, 2004; Ohlenschläger *et al.*, 2004; Ihle *et al.*, 2005; Melchers *et al.*,
2006). Этот пространственный структурный класс является одним из самых широко представленных в рибосомальной РНК (Woese *et al.*, 1990; Antao, Lai & Tinoco, 1991; Antao & Tinoco, 1992).

Таким образом, несмотря на значительный прогресс в изучении молекулярных аспектов инициации репликации генома вируса полиомиелита, принципы взаимного узнавания репликативного элемента *ori*L и вирусного белка 3CD остаются не ясными. Кроме того, остается непонятным, как реализуются эти принципы при высокой скорости накопления мутаций и в условиях бутылочного горлышка, которые характерны для энтеровирусов.

#### 1.3 Пространственные структурные классы тетрапетель

На данный момент известны несколько пространственных структурных классов тетрапетель. Три из них наиболее часто встречаются в рибосомальных РНК и являются наиболее изученными - это классы UNCG (N=любой нуклеотид), GNRA (R=A/G) и gCUUGc (Woese *et al.*, 1983, 1990; Antao *et al.*, 1991; Antao & Tinoco, 1992; Proctor *et al.*, 2002; Cheong & Cheong, 2010; Bottaro & Lindorff-Larsen, 2017)(Рисунок 12). Пространственная структура тетрапетель этих классов существенно различается и будет рассмотрена ниже.



Рисунок 12. Схематическое представление взаимодействий внутри тетрапетель трех пространственных структурных классов. Условные обозначения: пунктиром отмечены водородные связи, черными прямоугольниками - стекинг взаимодействия, пятиугольниками обозначены рибозы (рибозы в конформации 2'-эндо - черные, рибозы в конформации 3'-эндо - белые), фосфаты обозначены кругами, основания в конформации *anti* обозначены белыми прямоугольниками, в конформации *syn* – серым прямоугольником) (лекция "Principles of Nucleic Acids", Prof. Dr. Robert Haener). Показан ход сахарофосфатного остова (отмечен трубкой) и взаимной ориентации первого и четвертого нуклеотидов представителей тетрапетель пространственных классов UNCG (A, идентификаторы в базе данных PDB - 1TXS, 2EVY), cGNRAg (Б, идентификаторы в базе данных PDB - 12IG) и gCUUGc (B, идентификаторы в базе данных PDB - 12IG) и gCUUGc (B, идентификаторы в базе данных PDB - 12IG) (Melchers *et al.*, 2006).

Только среди представителей этих трех пространственных структурных классов встречаются тетрапетли с аномально высокой температурой плавления (Tm) (Proctor *et al.*, 2002). Для тетрапетель cUUCGg и cGAGAg в разных исследованиях температура плавления составляет около 70°C и выше (Antao & Tinoco, 1992; Varani, 1995; Sheehy, Davis & Znosko, 2010). Температура плавления сильно зависит от фланкирующей пары как для тетрапетель, принадлежащих консенсусам YNMG и GNRA, так и для всех остальных тетрапетель. Для большинства тетрапетель, кроме gCUUGc, прилегающие пары снижают температуру в следующем ряду: c-g<g-c<u-g (Antao *et al.*, 1991; Melchers *et al.*, 2006; Sheehy *et al.*, 2010). Например, у шпильки gccGAGAggc температура плавления 69.6°C, у шпильки gcuGAGAggc температура плавления ниже и составляет 48.6°C (Sheehy *et al.*, 2010). У шпильки

uacUUCGgua температура плавления 75.2°C, у шпильки uauUUCGgua температура плавления 59.7°C (Melchers *et al.*, 2006).

Существуют также и другие классы тетрапетель, не обладающих свойством экстремальной термостабильности: классы ANYA, AGNN и GANC (Correll & Swinger, 2003; Keating, Toor & Pyle, 2008) (Butcher, Dieckmann & Feigon, 1997; Lebars *et al.*, 2001; Wu *et al.*, 2001, 2004). Например, тетрапетля AUUA с фланкирующими парами u-g или c-g имеет температуру плавления в диапазоне от 53°C до 63°C, что примерно на десять градусов ниже, чем у представителей классов YNMG и GNRA с фланкирующей парой c-g (Serra *et al.*, 1997). Такие тетрапетли встречаются гораздо реже и не будут использованы в данной работе, поэтому на них мы не будет останавливаться подробно.

Следует отметить, что РНК имеет динамическую природу и не покоится постоянно в одной и той же конформации (Ferner *et al.*, 2008; Bothe *et al.*, 2011; Al-Hashimi, 2013; Bottaro, Di Palma & Bussi, 2014; Bottaro *et al.*, 2016a; Bottaro, Gil-Ley & Bussi, 2016b). Например, методом симуляции молекулярной динамики *in silico* показано, что тетрапетля GAAA, которая в большинстве исследований демонстрирует пространственную структуру класса GNRA, также способна формировать структуру пространственного класса UNCG (Butcher *et al.*, 1997; Bottaro *et al.*, 2016b).

#### 1.3.1 Пространственная структура тетрапетель класса UNCG

Класс тетрапетель yUNCGg характеризуется несколькими структурными особенностями (Varani, Cheong & Tinoco, 1991; Ennifar *et al.*, 2000; Du *et al.*, 2003, 2004). Основание четвертого нуклеотида находится в конформации «*syn*» к рибозе, остальные основания в конформации «*anti*» (Рисунок 12, 13, 14).

В результате первый и четвертый нуклеотиды формируют неканоническую *anti-syn* пару, связанную двумя водородными связями (между 2'-гидроксилом рибозы первого нуклеотида и кислородом основания четвертого нуклеотида и между кислородом основания первого и имино группой четвертого нуклеотидов). Основание третьего нуклеотида (А или С) образует стэкинг с основанием первого либо с основаниями первого и второго нуклеотидов петли и водородную связь с фосфатной группой сахарофосфатного остова между первым и вторым нуклеотидом. Основания первого и последнего нуклеотидов петли образуют стэкинг с основаниями фланкирующей пары петли. Таким образом, наблюдается продолжительный стэкинг, включающий прилегающие к тетрапетле фланкирующие пары, пару первого и четвертого нуклеотида тетрапетли и третий нуклеотид петли у некоторых представителей этого структурного класса (Рисунок 12). Основание второго нуклеотида не образует никаких связей с остальными нуклеотидами, свободно вращается

со стороны малого желобка и может вступать в различные взаимодействия (Рисунок 14 14).



Рисунок 13. (A) Конформация рибоз (3'-эндо и конверт 2'-эндо) и (Б) конформации нуклеотидов анти (anti) и син (syn) на примере пиримидинового нуклеотида; рибозы обозначены пятиугольниками, плоскость основания - прямоугольником. Основание поворачивается относительно неподвижного сахара. Торсионный угол O4'-C1'-N1-C2 обозначен χ. (Зенгер В. «Принципы организации нуклеиновых кислот» Москва; «Мир».1987)

Для шпилек пространственного консенсуса YNMG и пространственного структурного класса UNCG замена аденозина/цитидина (М) в третьем положении на гуанозин/уридин приводит к понижению температуры плавления, что объясняется отсутствием водородной связи основания третьего нуклеотида и фосфата сахарофосфатного остова (Proctor *et al.*, 2002). К тем же результатам приводит и замена фланкирующей пары C-G на G-C (Proctor *et al.*, 2002).

Показано, что тетрапетля GCUA с фланкирующей парой u-g относится к пространственному классу UNCG (Рисунок 12, 14) (Melchers *et al.*, 2006). Первый и четвертый нуклеотиды этой тетрапетли формируют неканоническую *«anti-syn»* пару (Рисунок 12). Второй нуклеотид обращен в малый желобок, основание третьего нуклеотида находится со стороны большого желобка и формирует стекинг с основанием первого нуклеотида. Отличиями от других тетрапетель пространственного класса UNCG являются 1) отсутствие водородной связи между основанием третьего нуклеотида и сахарофосфатным остовом, 2) преимущественное расположение основания второго нуклеотида в малом желобке (возможно формирование водородной связи между основаниями второго нуклеотида и нуклеотида фланкирующей пары), 3) наличие водородной связи между между рибозой первого нуклеотида и фосфатной группой между вторым и третьим нуклеотидом (Melchers *et al.*, 2006).



Рисунок 14. Схематическое представление пространственных структур тетрапетель uUACGg (A), uGCUAg (Б), cGAAAg (В) с фланкирующей парой со стороны малого желобка (по Melchers et al. 2006) (основания обозначены цветами: U обозначен оранжевым, G обозначен зеленым, A обозначен красным, C обозначен желтым).

В РНК разных организмов, в том числе вирусов, известны шпильки с пятью или с тремя нуклеотидами, которые имеют структурные признаки пространственного класса UNCG. Это тетрапетли cUCAGUg и uCUUGUa, в которых первые четыре нуклеотида формируют типичную пространственную структуру класса UNCG, а пятый выпетлен в большой желобок (Theimer, Finger & Feigon, 2003; Liu *et al.*, 2009). Еще одним представителем является тринуклеотидная петля домена d *oriL* риновируса 14, нуклеотиды которой имеют конформацию, аналогичную первому, третьему и четвертому нуклеотидам пространственного класса UNCG (Headey *et al.*, 2007).

#### 1.3.2 Пространственная структура тетрапетель класса GNRA

Тетрапетли пространственного структурного класса GNRA – вторые по распространенности в рибосомальной PHK после тетрапетель пространственного структурного класса UNCG (Woese *et al.*, 1990). На сегодняшний день структура GNRA продемонстрирована для следующих тетрапетель: GAAA, GCAA, GAGA, GCGA, GUAA (Heus & Pardi, 1991; SantaLucia, Kierzek & Turner, 1992; Cate *et al.*, 1996; Correll, Wool & Munishkin, 1999; Mohammed *et al.*, 2014). Впервые структура GNRA была определена Heus и Pardi (1991). Было показано, что все основания тетрапетли повернуты в сторону малого желобка и находятся в ориентации *anti* относительно рибоз (Рисунок 12). Первое и четвертое основания образуют в некоторых случаях водородные связи и формируют неканоническую пару (Mohan *et al.*, 2010). Основание третьего нуклеотида образует стекинг с четвертым. Основание второго нуклеотида также может формировать стекинг с основанием третьего нуклеотида, а может свободно вращаться со стороны малого желобка (Heus & Pardi, 1991) (Рисунок 12, 14).

К пространственному классу GNRA по взаимной ориентации элементов пространственной

структуры можно отнести тетрапетли консенсуса UNAC (Zhao *et al.*, 2012). Известны пентапетли, четыре первых нуклеотида которых ориентированы аналогично четырем нуклеотидам тетрапетель пространственного класса GNRA, а пятый нуклеотид выпетлен (Cai *et al.*, 1998; Schärpf *et al.*, 2000).

#### 1.3.3 Пространственная структура тетрапетель класса gCUUGc

Пространственный структурный класс gCUUGc назван по представителю класса - тетрапетле gCUUGc (Jucker & Pardi, 1995). В этой тетрапетле основание второго нуклеотида завернуто в малый желобок и формирует там водородные связи с парой оснований первого и четвертого нуклеотидов и парой, прилегающей к тетрапетле (Рисунок 12). Основания первого и четвертого нуклеотида формируют классическую Уотсон-Криковскую пару с ориентацией оснований относительно рибоз *anti-anti*. Основание третьего нуклеотида образует стекинг с основаниями первого и четвертого нуклеотого нуклеотидов примерно в половине моделей, сформированных по данным ЯМР (Jucker & Pardi, 1995).

К классу gCUUGc также относят пентапетли uCUGGCa и gCAGGUc (Oberstrass *et al.*, 2006; Schwalbe *et al.*, 2008). В обоих случаях пятый нуклеотид петли выпетлен и не оказывает влияния на ориентацию остальных четырех нуклеотидов пентапетли. В случае uCUGGCa второй нуклеотид стабилизирован в малом желобке водородными связями, основания первого и четвертого нуклеотидов образуют Уотсон-Криковскую пару (Oberstrass *et al.*, 2006). В случае gCAGGUc основания второго и третьего нуклеотидов завернуты в малый желобок, первый и четвертый нуклеотиды формируют Уотсон-Криковскую пару (Schwalbe *et al.*, 2008). Несмотря на сильное различие между структурами gCAGGUc и gCUUGc (третий нуклеотид gCAGGUc завернут в малый желобок и не образует стекинг с основаниями первого и четвертого нуклеотидов, как в gCUUGc), авторы относят gCAGGUc к пространственному классу gCUUGc на основании схожести остальных параметров (Schwalbe *et al.*, 2008).

#### 2. Материалы и методы

#### 2.1 Реагенты

В работе были использованы ферменты, произведенные фирмами Fermentas (Вильнюс, Литва), Promega (Фитчбург, США), Invitrogen (Карлсбад, США), Qiagen (Венло, Нидерланды).

Были использованы реагенты и наборы реагентов, произведенные фирмами Calboichem, Pharmacia, Qiagen, Fermentas, Merk, Serva, Sigma, GE Healthcare.

В работе были использованы олигонуклеотиды производства компании Syntol (Таблица 2).

Таблица 2 Нуклеотидные последовательности использованных в работе олигонуклеотидов и их соответствие последовательности генома полиовируса.

N	Название	Последовательность <sup>а</sup>	Позиция в геноме
1	SPLI	CTTGGTTTcGTaCGTCTAAG	5-124, комплемент
2	Rib2	gaggccgaaaggccgaaaagggcctatgggcccttcTTAAAACAGCTCTGG	1-15
3	BHT7Rib1	Ctggatcctaatacgactcactatagggtgttttaactgatgaggccgaaaggccg	нетб
4	SolD	gacgggcccttcTTAAAACAGCTCTGGGGTTGTACCCACCCCAGA GGCCCACGTGGCGGCTAGTACTCCGGT(N8)ACCCTTGTACG CCTGTTTTATACTCCCTTCCCGTAACTTAGACGTACGAA	1-118
5	mut-s	CTCCGGT(mut-s)ACCCTTG <sup>B</sup>	54-75
6	mut-a	GT(mut-a)ACCGGAGTAC <sup>B</sup>	51-70, комплемент
7	3EP4	TTTTTTTTTTTTTTCTCCG	7438-polyA
8	DEN3	GAAACAGAAGTGCTTGTTCG	158-177, комплемент
9	H40As	ACCAACCgcCGCTTCACCTG	5548-5567
10	H40Aa	GAAGCGgcGGTTGGTAAAATAG	5541-5562, комплемент
11	3D3'	tataagatctAAATGAGTCAAG	7358-7369, комплемент
12	3CD5'	gccatgGGACCAGGGTTCGATTAC	5438-5458
13	B5594	GAAGTGGAGATCTTGGATGCC	5594-5614
14	DP7	TTCCTTCGAAGGTCTCATCC	5987-6019, комплемент
15	PVL1	GGCAGACGAGAAATACCCAT	7122-7141

16	PVR1	CGAACGTGATCCTGAGTGT	7210-7229, комплемент
17	PVP1	FAM-GTTGATTCATGAATTTCCTTCATTGGCA-BHQ1 <sup>r</sup>	7160-7186

<sup>а</sup> строчными буквами обозначены измененные или дополнительные нуклеотиды, заглавными буквами обозначены нуклеотиды, соответствующие геному вируса полиомиелита

<sup>б</sup>последовательность вектора pT7PV1Rib(+)MS

<sup>в</sup> олигонуклеотиды вместо (mut-s) и (mut-a) содержат последовательности измененного октануклеотида и его комплемента, соответственно

<sup>г</sup> олигонуклеотид на 5'конце комплементарно соединен с флуоресцентным красителем карбоксифлуоресцеином (FAM) и на 3'конце комплементарно соединен с гасителем Black Hole Quencher-1 (BHQ1)

#### 2.2 Штаммы E.coli и получение компетентных клеток

#### 2.2.1 Штамм ТОР10

Источник штамма: Invitrogen

Генотип: F<sup>-</sup> mcrA  $\Delta$ (mrr-hsdRMS-mcrBC)  $\varphi$ 80lacZ $\Delta$ M15  $\Delta$ lacX74 nupG recA1 araD139  $\Delta$ (ara-leu)7697 galE15 galK16 rpsL(StrR) endA1  $\lambda$ -

Штамм лишен плазмиды F (F<sup>-</sup>), ответственной за обмен генетическим материалом между бактериями, что позволяет сохранять постоянным генотип штамма и целостность плазмид, которыми он трансформирован. Также штамм содержит мутации в генах эндонуклеаз (делеция  $\Delta$  (mrr-hsdRMS-mcrBC, мутация в гене неспецифической эндонуклеазы 1 - endA1)) и гене, ответственном за рекомбинацию (мутация гесA1), что обеспечивает сохранение последовательности целевой плазмиды. Штамм не способен метаболизировать арабинозу, галактозу и лейцин (мутация araD139, делеция  $\Delta$ (ara-leu)7697 и мутация galK16). Гены, ответственные за синтез дезоксирибозы, экспрессируются конститутивно (мутация nupG), что обеспечивает возможность трансформации штамма плазмидами большого размера. Мутация galK16 обеспечивает повышенную эффективность трансформации благодаря снижению количества липосахаридов в клеточной стенке. Также штамм позволяет осуществлять скрининг колоний с помощью так называемого бело-голубого теста, так как имеет делецию в N-концевом участке гена β-галактозидазы и делецию Lac-оперона (делеции  $\phi$ 80lacZ $\Delta$ M15 и  $\Delta$ lacX74). Штамм TOP10 устойчив к антибиотику стрептомицину (мутация rpsL(StrR)) и не подвержен лизогении под действием фага лямбда ( $\lambda$ -).

#### 2.2.2 Штамм ЈМ109

Источник штамма: Promega.

Генотип: endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB<sup>+</sup> $\Delta$ (lac-proAB) e14- [F' traD36 proAB<sup>+</sup>lacI<sup>q</sup>lacZ $\Delta$ M15] hsdR17(r<sub>K</sub><sup>-</sup>m<sub>K</sub><sup>+</sup>)

Штамм JM109 имеет мутации в гене эндонуклеазы 1 (endA1) и в гене, ответственном за рекомбинацию (recA1). Дополнительно штамм JM109 имеет мутацию glnV44, приводящую к супрессии амбер-мутации; чувствителен к недостатку тиамина (мутация thi-1); устойчив к нафтиридон-карбоновой кислоте (мутация в гене гиразы gyrA96); сохранил рестриктазную активность относительно метилированного сайта RmC (mycB+); не имеет активности рестриктазы EcoKI, но способен метилировань ДНК по ее сайту (hsdR17(rK-mK+)); содержит мутацию в Laconероне и генах биосинтеза пролина ( $\Delta$ (lac-proAB)); не несет профага e14 (e14-) и поэтому не имеет эндонуклеазной активности относительно метилированного сайта CmCGG. Кроме того, штамм содержит плазмиду F, которая несет в себе мутацию, инактивирующую ее трансфер (traD36), гены биосинтеза пролина (proAB+), делецию в N-концевом участке  $\beta$ -галактозидазы (lacZ $\Delta$ M15, позволяет осуществлять бело-голубой тест) и мутацию lacIq в промоторе ингибитора лактозного оператора (lacI), приводящую к увеличению его экспрессии в десять раз по сравнению с уровнем экспрессии у бактерий дикого типа. Последняя мутация позволяет эффективно контролировать экспрессию белков с плазмид, содержащих LacO, что значимо при клонировании токсичных для *E.coli* белков.

#### 2.2.3 Получение компетентных клеток E.coli

Бактериальные клетки высевали штрихом на чашку с агаризованной средой SOB (Super Optimal Broth) (1% Васtotryptone, 0,5% дрожжевого экстракта, 10 mM NaCl, pH 7,5). На следующий день отдельную колонию высевали в 3 мл среды SOB (2% бактотриптон, 0,5% дрожжевого экстракта, 10 mM NaCl, 2,5 mM KCl, pH 7,5) и растили ночь при  $37^{\circ}$ С, непрерывно встряхивая в ротационной качалке. 250 мкл ночной культуры пересевали в 50 мл среды SOB и инкубировали на качалке при  $37^{\circ}$ С до тех пор, пока титр клеток не достигал  $4-7*10^7$  КОЕ/мл. Для этого определяли оптическую плотность при 600 нм. Нужному титру соответствует оптическая плотность 0,5. Клетки осаждали центрифугированием в стерильных пробирках при 3500g в течение 15 мин. Осадок тщательно ресуспендировали стерильной пипеткой в 22,5 мл буферного раствора RF1 (100 mM RbCl<sub>2</sub>, 50 mM MnCl<sub>2</sub>, 30 mM CH<sub>3</sub>COOK, 10 mM CaCl<sub>2</sub>, 15% глицерин, pH 5,8) и выдерживали 120 мин в ледяной бане. После этого клетки снова собирали центрифугированием в

том же режиме, ресуспендировали в 4 мл буферного раствора RF2 (10 mM MOPS (3-(N-morpholino)propanesulfonic acid), 10 mM RbCl2, 75 mM CaCl2, 15% глицерин, pH 6,8) и инкубировали 30 мин в ледяной бане. Суспензию клеток расфасовывали по 100 мкл в стерильные пробирки и хранили при  $-80^{\circ}$ C.

#### 2.3 Плазмиды и базовые генно-инженерные методы

#### 2.3.1 Плазмида pT7PV1(Rib+)MS

В работе использовались следующие плазмиды: **pT7PV1(Rib+)MS** – сконструирована на основе плазмиды pT7PV1 (Pilipenko *et al.*, 1992а). Схема плазмиды представлена на рисунке (Рисунок 15). Эта плазмида кодирует полный геном вируса полиомиелита под контролем промотера T7. Последовательность вирусного генома предваряется последовательностью рибозима типа Hammerhead (см. Рисунок 15, Рисунок 16). В молекуле PHK, полученной в результате транскрипции с этой плазмиды, рибозим отщепляется, оставляя 5' конец полностью соответствующим нуклеотидной последовательности 5'-конца генома вируса полиомиелита (Herold & Andino, 2000) (Рисунок 16). Показано, что такая модификация приводит к более раннему появлению вирусных бляшек и увеличению удельной инфекционности транскрипта не более чем в два раза.



Рисунок 15. Схема строения плазмиды pT7PV1Rib(+)MS.



Рисунок 16. Вторичная структура и последовательность рибозима типа Hammerhead на 5'конце полногеномного транскрипта. Стрелкой указано место расщепления.

# 2.3.2 Конструирование плазмиды pT7PV1Rib(+)MS с рандомизированным октануклеотидом в домене d



Рисунок 17 Схема получения плазмиды pT7PV1(Rib+)MS с рандомизированным участком

Для получения плазмиды pT7PV1Rib(+)MS с рандомизированным октануклеотидом в домене d oriL было осуществлено лигирование двух фрагментов: фрагмента pT7PV1Rib(+)MS ApaI-SplI плазмиды И фрагмента, имеющего рандомизированный участок (Prostova et al., 2015) (Рисунок 17). Последний был получен с помощью ПЦР на матрице олигонуклеотида SolD (10 пкмоль) (см. Таблицу 1) с праймерами SPLI и Rib2 (по 60 пкмоль). Количество матрицы (6,02×10<sup>12</sup>) было подобрано так, чтобы достоверно перекрыть количество возможных сочетаний в рандомизированном октануклеотиде (65536 шт.). ПЦР осуществлялся в следующих условиях: 55°C – 20 сек, 72 °C - 20 сек, 95 °C - 15 сек - всего 10 циклов. Продукт ПЦР и плазмида были обработаны эндонуклеазами ApaI и SplI, и полученные фрагменты нужной длины были использованы для реакции лигирования.

#### 2.3.3 Конструирование плазмид pT7PV1Rib(+)MS с измененными oriL

Для получения плазмид pT7PV1Rib(+)MS, несущих заданные мутации, было осуществлено лигирование двух фрагментов. Первый фрагмент ApaI-SplI (длинный фрагмент) был получен путем рестрикции плазмиды pT7PV1Rib(+)MS. Второй фрагмент, несущий заданную мутацию, был получен с помощью одного из двух вариантов ПЦР (Рисунок 15). В случае реконструкции мутаций, обнаруженных в вирусной последовательности, ПЦР осуществлялся на матрице кДНК, полученной в реакции обратной транскрипции с тотальной РНК из инфицированной вирусом культуры клеток с праймером 3EP4 или шестичленным рандомным олигонуклеотидом. В случае если необходимые мутации вводились *de novo*, был использован двустадийный фьюжн-ПЦР с праймерами SPLI, Rib1, прямым и обратным праймерами для мутагенеза mut-s и mut-a на матрице плазмиды pT7PV1Rib(+)MS (Таблица 2).

#### 2.3.4 Плазмида pQE60

Плазмида pQE60 производится компанией Qiagen. Схема плазмиды представлена на рисунке Рисунок 18). Плазмида содержит промотор фага T5 (узнается PHK-полимеразой I *E.coli*), сайт начала репликации ColE1, ген бета-лактамазы (обуславливает устойчивость к ампициллину), и два Lacоператора (LacO), обеспечивающие присоединение репрессора LacI для контроля экспрессии.



Рисунок 18. Схема строения плазмиды pQE60. Отмечены промотер T5, Lac – опероны, сайты рестрикции, сайта начала репликации Col E1 (материалы Qiagen).

#### 2.3.5 Конструирование плазмиды, экспрессирующей вирусный белок 3CD

Для клонирования и экспрессии вирусного белка 3CD были опробованы четыре вектора (pTYB21 (NEB), pTrcHis A (Invitrogen), pQE60 (Qiagen), pET15b (Novagen)) и несколько штаммов E.coli, в зависимости от используемого в векторе промотора (JM109, BL21, BL21 rosetta, HB101, DHa). Удачным сочетанием для клонирования и экспрессии 3CD оказались вектор pQE60 и штамм *E.coli* JM109.



Рисунок 19 Схема клонирования плазмиды pQE60-3CD

Последовательность 3CD была клонирована в вектор pQE60 (Qiagen) в два этапа (Рисунок 19).

На первом этапе в вектор была осуществлена вставка участка генома вируса полиомиелита по сайтам NcoI и BgIII, кодирующего белок 3CD (нуклеотиды 5432 по 5606), содержащего H40A аминокислотную замену в протеолитическом сайте, для того, чтобы избежать протеолиза белка 3CD при хранении (Рисунок 19, 1). Необходимый фрагмент был получен с помощью двухстадийного фьюжн-ПЦР с праймерами 3CD5', DP7 и мутагенными праймерами H40Aa и H40As на матрице плазмиды pT7PV1Rib(+)MS (Таблица 2). Вторым этапом вводили оставшуюся последовательности, часть кодирующей 3CD (нуклеотиды с 5607 по 7369) с использованием сайта рестрикции BglII (Рисунок 19, 2). Фрагмент для второго этапа был получен с помощью ПЦР с праймерами B5594 и 3CD3' на матрице плазмиды pT7PV1Rib(+)MS. Получить правильный клон оказалось возможным только при высеве трансформированных лигазной смесью клеток на чашку со средой, содержащей 20 mM глюкозу. Наличие глюкозы дополнительно подавляло экспрессию 3CD под контролем LacO. Правильную

ориентацию вставки подтверждали аналитической рестрикцией с помощью ферментов рестрикции SacII и NcoI. Было проанализировано 20 клонов и отобран один, содержащий правильную

ориентацию вставки. Однако этот клон имел делецию 20 нуклеотидов, захватывающую последовательность Шайна-Дальгарно плазмиды. Последовательность Шайна-Дальгарно была восстановлена путем генно-инженерного переноса участка NcoI-XhoI из плазмиды PQE60 в плазмиду pQE60-3CD. Отсутствие дополнительных мутаций подтверждали секвенированием.

#### 2.3.6 Трансформация компетентной культуры клеток плазмидной ДНК

К 100 мкл компетентных клеток добавляли раствор ДНК (до 1 мкг в объёме до 25 мкл), тщательно перемешивали стерильным наконечником и оставляли в ледяной бане на 30 мин Затем пробирку переносили в водяную баню или термостат с температурой 42°C на 2 мин, после чего охлаждали в ледяной бане в течение 3 мин, добавляли 400 мкл среды SOB и инкубировали в термостате при 37°C 1 час. После этого бактерии высевали на чашки с агаризованной средой SOB, содержащей 100 мкг/мл ампициллина.

#### 2.3.7 Выделение плазмидной ДНК

Для выделения плазмидной ДНК применяли метод щелочного лизиса. Отдельные колонии высевали в 3 мл среды SOB, содержащей 100 мкг/мл ампициллина, растили ночь при 37°C в ротационной качалке. Выросшую культуру переносили в пробирки типа Eppendorf, и клетки осаждали центрифугированием при 12000 об./мин в течение 3 мин Осадок промывал 500 мкл буферного раствора TNE (10 mM Tris-HCl, pH 8,0, 100 mM NaCl, 1 mM EDTA), суспендировали в 100 мкл раствора I (50 mM глюкоза, 25 mM Tris-HCl, pH 8,0, 10 mM EDTA) с 5 мг/мл лизоцима (лизоцим растворяли в растворе I непосредственно перед использованием) и оставляли при комнатной температуре на 30 мин Затем в каждую пробу добавляли по 200 мкл свежеприготовленного раствора II (0,2 M NaOH, 1% SDS) и аккуратно перемешивали. После 10 минной инкубации добавляли 150 мкл раствора, содержавшего 3 М К<sup>+</sup> и 5 М СН<sub>3</sub>СОО<sup>-</sup>, энергично встряхивали и выдерживали 40 мин при 4°С. После этого пробы центрифугировали в течение 15 мин в том же режиме и отбирали супернатант. Нуклеиновые кислоты из супернатанта экстрагировали один раз 500 мкл смесью фенола, насыщенного TNE, с хлороформом (1:1) и осаждали добавлением 1 мл этанола в течение 30 мин при -70°С. Осадок собирали центрифугированием в том же режиме в течение 15 мин, промывали 80%-ным этанолом и высушивали в вакуумном эксикаторе. Сухой осадок растворяли в 100 мкл воды, добавляли 3 мкл РНКазы А (20 мг/мл) и инкубировали 1 час при 37°С. ДНК осаждали добавлением 80 мкл раствора, содержащего 20% полиэтиленгликоля и 2,5 М NaCl, в течение 1 часа при 4°C. Нуклеиновые кислоты осаждали центрифугированием в том же режиме в течение 15 мин, растворяли в 100 мкл воды, дважды экстрагировали 100 мкл смесью фенола, насыщенного TNE, с хлороформом (1:1) и осаждали путём добавления 80 мкл 5 М ацетата аммония и 600 мкл этанола в течение 1 часа при -70°С или ночи при -20°С. Для дальнейшей работы плазмиду осаждали на микроцентрифуге (15 мин), промывали 80%-ным этанолом, высушивали в вакуумном эксикаторе и растворяли в требуемом объёме воды.

Также для выделения плазмидной ДНК суспензии *E.coli* использовали набор для выделения плазмидной ДНК фирмы Fermentas «GeneJET Plasmid Miniprep Kit», согласно протоколу производителя.

Для выделения нуклеиновых кислот из большого количества суспензии *E.coli* объемы всех растворов увеличивали соответственно.

#### 2.3.8 Обработка плазмидной ДНК эндонуклеазами

К водному раствору ДНК добавляли 1/10 объема соответствующего коммерческого 10кратного буферного раствора, 1-100 ед. эндонуклеазы (в зависимости от количества ДНК) и инкубировали от 1 до 3 часов при температуре 37°С. После окончания инкубации для экстракции плазмидной ДНК добавляли 0,5 объема насыщенного фенола и 0,5 объема хлороформа. Фенол предварительно насыщали буферный растворным раствором TNE (10 mM Tris pH8, 100 mM NaCl, 0,1 mM EDTA) до значения pH>7. Смесь интенсивно перемешивали в течение 5-10 мин и центрифугировали 5 мин при 13000 оборотов/мин. Супернатант отбирали и осаждали ДНК добавлением 0,8 объема 5М ацетата аммония и 2.5 объемов 96%-ного этилового спирта или 0,5 объема 5 М ацетата аммония и 2 объемов изопропилового спирта. Осадок промывали 80%-ным этиловым спиртом, высушивали и растворяли в воде.

#### 2.3.9 Лигирование ДНК

Для сборки кольцевых плазмид из линейных молекул ДНК нужные фрагменты смешивали в эквимолярном соотношении, добавляли 2 мкл 10-кратного коммерческого буферного раствора, 5 ед. ДНК-лигазы фага Т4 и объем смеси доводили водой до 20 мкл. Смесь инкубировали в течение ночи при 12-14°C. На следующий день добавляли еще столько же ДНК-лигазы и инкубировали 1 час при комнатной температуре.

#### 2.3.10 Аналитический электрофорез в агарозном геле

Использовали 1%-ный агарозный гель. К 1 г порошка агарозы (Sigma type I) добавляли 10 мл 10-кратного буферного раствора ТВЕ (890 mM Tris, 890 mM борная кислота, 20 mM ЭДТА) и доводили объем водой до 100 мл. Агарозу расплавляли на кипящей водяной бане и добавляли 8 мкл насыщенного раствора бромистого этидия. Гель заливали в плашку с гребенкой и оставляли застывать при комнатной температуре. Раствор ДНК перед нанесением на гель смешивали с <sup>1</sup>/<sub>4</sub> объема буферного раствора для нанесения (5-кратный ТВЕ, 30% глицерина, 0,25% бромфенолового синего). Электрофорез проводили в буферном растворе ТВЕ при напряженности 5-7,5 В/см. Для электрофореза РНК в ТВЕ добавляли SDS до 0,01%.

### 2.4 Транскрипция и очистка РНК

#### 2.4.1 Транскрипция in vitro

Полногеномную РНК получали с использованием ДНК-зависимой РНК-полимеразы фага T7 (Fermentas/Thermo Scientific). В качестве матрицы использовали линеаризованные плазмиды pT7PV1Rib(+)MS с введенными мутациями. В случае получения полногеномного транскрипта для линеаризации использовали фермент EcoRI (ThermoScientific). Для получения *oriL* плазмида pT7PV1(+)MS была линеаризована ферментом SplI (или ее изошизомером Pfl23II) (Fermentas). Транскрипцию осуществляли в буферном растворе 20 mM Tris-HCl pH 7.8, 5 mM MgCl<sub>2</sub>, 1 mM Спермидин (SERVA) в присутствии ингибитора PHKa3 Ribolock (Fermentas/Thermo Scientific) или RNasin (Promega) 1.2 ед./мкл, также добавляли DTT (Amersham) до 10 mM, смесь рибонуклеотидов до 2 мМ каждого и T7 полимеразу до концентрации 1 ед./мкл. Для получения полногеномного транскрипта реакцию инкубировали 3 часа при  $37^{\circ}$ С. Для получения *oriL* реакцию осуществляли в течение 2-х часов при  $37^{\circ}$ С, затем добавляли раствор MgCl<sub>2</sub> до концентрации 20 мМ и инкубировали еще 1 час при  $57^{\circ}$ С.

В обоих случаях реакцию останавливали добавлением 4 М EDTA до концентрации 20 mM.

#### 2.4.2 Очистка полногеномных транскриптов в сахарозном градиенте

Полногеномные транскрипты очищали от примеси ДНК и фрагментов РНК методом центрифугирования в градиенте сахарозы 5-20%, растворенной в буферном растворе (10 mMTris-HCl pH 7,5; 150 mM NaCl; 1 mM EDTA; 0,1% SDS). Градиент готовили в центрифужных пробирках и наслаивали на него реакционную смесь, разведенную в два раза двукратным буферным раствором (20 mM Tris-HCl pH 7,5; 300 mM NaCl; 2 mM EDTA; 0,2% SDS). Далее осуществляли центрифугирование при 40 тыс. оборотов/мин в роторе SW40, центрифуге L7-65 (Beckman) при 16°C в течение 4 ч. По окончании центрифугирования градиент раскапывали на аликвоты по 750 мкл. Аликвоты, содержащие полногеномный транскрипт, отбирали по результатам электрофореза в 1%ном агарозном геле. РНК осаждали дважды добавлением 0,8 объема 5 М ацетата аммония и 2,5 объемов 96% этанола. Осадок промывали 80%-ным этанолом, высушивали и растворяли в воде. Концентрацию и чистоту РНК в пробе определяли спектрофотометрически.

#### 2.4.3 Получение и очистка oriL

Для получения транскрипта, соответствующего первым 113 нуклеотидам полиовирусного генома, линеаризованную по сайту рестрикции Spll плазмиду pT7PV1MS использовали в качестве матрицы для транскрипции in vitro. По окончании реакции транскрипции ДНК-мартрицы к реакционной смеси добавляли сухую мочевину до 6 М, буферный раствор для нанесения (50% глицерин, 0,01% бромфеноловый синий) и наносили после проведения префореза (1 ч, 200 V) на 12.5%-ный ПААГ, содержащий 6 М мочевину и 2-кратный буферный раствор ТАЕ. Электрофорез проводили при 100 V без охлаждения. По окончании электрофореза край геля погружали в краску с толуидином (уксусная кислота 10%, этанол 10%, толуидин 0,7%) для визуализации РНК. По результатам окрашивания из геля стерильным шпателем вырезали полосу, соответствующую по длине oriL без рибозима (средняя полоска из трех). Гель разминали между двумя листами парафильма (SERVA) и помещали в пробирку с буферным растворным раствором Tris-HCl 50 mM pH7.5, NaCl 100 mM, EDTA 2 mM, SDS 0,1%. Элюировали РНК из геля перемешиванием в течение 1-2 часов при +4°С. Затем центрифугировали при 3000 оборотов/мин при +4°С. Оставшийся гель еще раз перемешивали с буферный растворным раствором и отжимали шприцом через фильтр с размером пор 0,4 мкм. Супернатант и фильтрат объединяли и наносили на колонку с 5 мл DEAEsepharose (Pharmacia), уравновешенную тем же буферный растворным раствором без SDS. Промывали двумя объемами того же буферного раствора без SDS. Элюировали тем же буферный растворным раствором без SDS, содержащим 1 М NaCl. Фракции, с пиковыми значениями оптической плотности при 280 нм, концентрировали бутанолом, обрабатывали равным объемом хлороформа и осаждали РНК изопропанолом (0,5 объема пробы 5 М ацетат аммония, 2 объема пробы изопропанола, инкубировали 10 мин при комнатной температуре, центрифугировали 15 мин при 13000 оборотов/мин на настольной центрифуге при +4°С). Осадок промывали 80% этанолом, высушивали и растворяли в воде. Концентрацию и чистоту РНК в пробе определяли спектрофотометрически.

#### 2.5 Культуры клеток и вирусологические методы

#### 2.5.1 Культуры перевиваемых клеток и их ведение

В работе использовалась линия перевиваемых эукариотических клеток *RD* (клетки рабдомиосаркомы человека) и линия перевиваемых эукариотических клеток *Vero* (клетки почки зеленой мартышки).

Культуру вели следующим образом: из флакона с трехсуточной культурой удаляли среду и обмывали два раза раствором Версена (производства ФГБНУ ИПВЭ им. М.П. Чумакова). Затем культуру инкубировали смесью 1:1 растворов Версен и раствора трипсина (производства ФГБНУ ИПВЭ им. М.П. Чумакова) при комнатной температуре 2-5 мин. После инкубации смесь удаляли и добавляли необходимое количество среды ДМЕМ (производства ФГБНУ ИПВЭ им. М.П. Чумакова) с 5% эмбриональной телячьей сыворотки и одну пятую часть объема переносили в новый флакон/флаконы. Флаконы инкубировали при температуре 37°С. Если культуру рассаживали на панели, то инкубацию проводили в инкубаторе с содержанием CO<sub>2</sub> 5%.

#### 2.5.2 Метод бляшек

Бляшки – это области лизиса на монослое культуры клеток, появляющееся в результате цитопатогенного действия вируса. Для получения бляшек вируссодержащую среду в нескольких десятикратных разведениях вносили на полный монослой перевиваемой культуры клеток *Vero*. Во флакон вносили 200 мкл вируссодержащей жидкости, в лунку шестилуночной панели (Corning) вносили 100 мкл вируссодержащей жидкости. Адсорбцию вируса проводили в течение 30 мин при комнатной температуре. В ходе адсорбции флаконы или панели периодически покачивали для равномерного распределения суспензии. Затем во флакон или лунку вносили покрытие, состоящее из равных объемов двух смесей: 1) H<sub>2</sub>O 65%, раствор Эрла 20%, бикарбонат натрия 6%, нормализованная бычья сыворотка (или эмбриональная сыворотка) 4%, прижизненный краситель нейтральный красный 4%, канамицин 1% и 2) 1.8% Бактоагар (Difco) в воде. Перед смешиванием агар плавили на кипящей водяной бане в течение 30 мин. После застывания покрытия флаконы или панели убирали в термостат с 37°С с подачей углекислого газа до 5%.

#### 2.5.3 Определение титра вируса

Для определения титра вируса в вирусной суспензии делали несколько десятикартных разведений в среде ДМЕМ (производства ФГБНУ ИПВЭ им. М.П. Чумакова) без сыворотки. Культуру клеток *Vero* в 6-ти луночных панелях либо во флаконах с предварительно удаленной ростовой средой инфицировали каждым из разведений в двух повторностях (50 мкл на лунку в

панели, либо 100 мкл на флакон). Проводили адсорбцию в течение 30 мин при комнатной температуре и заливали покрытием, как описано выше. Через трое суток инкубации при 37°C с подачей углекислого газа до 5% проводили подсчет вирусных бляшек для каждого разведения. Для определения титра вируса использовали то разведение, при котором во флаконах или лунках вырастало 10 до 50 бляшек.

#### 2.5.4 Размножение вируса

Из флакона или панели с культурой отбирали среду и добавляли суспензию вируса. Культуру клеток с вирусом инкубировали 30 мин при комнатной температуре. Затем добавляли культуральную среду без сыворотки. Флаконы или панели инкубировали при 37°C с подачей углекислого газа до 5% (панели) или без подачи углекислого газа (флаконы) до появления тотального цитопатического эффекта (определяли с помощью светового микроскопа). При появлении цитопатического эффекта флаконы замораживали.

#### 2.5.5 Трансфекция культуры эукариотических клеток.

Для трансфекции использовали культуру клеток *Vero*. Водный раствор полногеномного транскрипта в нужных концентрациях смешивали в пропорции 1 к 10 со смесью 300 мкг/мл декстран 6000, PBS (137 mM NaCl, 2.7 mM KCL, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH7.4). Смесь инкубировали при +4°C 30 мин. Перед внесением PHK культуру клеток дважды промывали поддерживающей средой. Затем к культуре клеток добавляли смесь с PHK и инкубировали 30 мин при комнатной температуре. После адсорбции, в зависимости от цели трансфекции, культуру или отмывали от остаточной PHK и добавляли поддерживающую среду, или заливали агаровым покрытием.

## 2.6 Определение последовательности вирусного генома и кинетики синтеза вирусной РНК

#### 2.6.1 Выделение тотальной РНК из инфицированной культуры клеток

Выделение тотальной РНК осуществляли с помощью pearentroв Trizol (LifeTechnoligies) или Tri-reagent (Sigma). На культуру клеток без среды наливали pearent (250 мкл для шестилуночных панелей, 125 мкл для 12-луночных панелей, 100 мкл для 24-луночных панелей). Затем клеточный лизат отбирали в пробирку и добавляли хлороформ (пятую часть от объема добавленного pearentra).

Трясли и инкубировали 10 мин при комнатной температуре. Центрифугировали 10 мин при 12000 оборотов/мин на настольной центрифуге. К верхней фракции добавляли перегнанный изопропанол (половину объема добавленного реагента). Инкубировали 10 мин при комнатной температуре и центрифугировали при тех же условиях. Осадок промывали 80%-ным этанолом, высушивали при комнатной температуре около 20 мин и растворяли в 30 мкл воды.

#### 2.6.2 Выделение вирусной РНК

Выделение вирусной РНК осуществляли двумя методами.

Первый метод состоит из трех последовательных экстракций фенолом и смесью фенола и хлороформа. На первой стадии производили экстракцию равным объемом фенола с добавлением SDS до 0,1%. На двух последующих стадиях экстракцию осуществляли смесью 0,5 объема пробы фенола и 0,5 объема пробы хлороформа. Из водной фазы нуклеиновую кислоту осаждали добавлением 0,8 объема пробы 5 М ацетата аммония и 2,5 объема пробы 96%-ного перегнанного этанола.

Второй метод использовали для выделения вирусной РНК из небольших объемов вирусной суспензии. Для выделения использовали реактивы и колонки из набора для выделения РНК (RNeasy MiniKit - Qiagen) согласно рекомендациям производителя.

#### 2.6.3 Обратная транскрипция in vitro

Для получения кДНК копии вирусной РНК использовали обратную транскриптазу Maxima (Fermentas) или SSII (LifeTechlogies), или M-MVL (Promega). Реакцию обратной транскрипции проводили согласно предлагаемым производителем фермента протоколам с использованием рандомного шестичленного олигонуклеотида или праймера 3EP4 (Таблица 2). В случае использования рандомизированного праймера использовали следующий протокол: 1) смешивали РНК (до 1 мкг на реакцию), нуклеотиды (до конечной концентрации 0,25 mM) и праймер, 2) инкубировали смесь 5 мин при  $65^{\circ}$ С, 3) охлаждали смесь в бане со льдом, 4) добавляли коммерческий буферный растворный раствор и обратную транскриптазу согласно указаниям производителя, и инкубировали смесь 1 ч при  $42^{\circ}$ С 5) для остановки реакции и разрушения остаточной РНК прогревали пробы при  $75^{\circ}$ С 10 мин. В случае использования специфического праймера использовали следующий протокол: 1) смешивали РНК (до 1 мкг на реакцию), нуклеотиды (до конечной концентрации 0,25 мМ) и праймер (до конечной РНК прогревали смесь 1 ч при  $42^{\circ}$ С 5) для остановки реакцию), нуклеотиды (до конечной концентрации 0,25 мМ) и праймер (до конечной концентрации 0,25 мМ) и праймер, 2) инкубировали следующий протокол: 1) смешивали РНК (до 1 мкг на реакцию), нуклеотиды (до конечной концентрации 0,25 мМ) и праймер, 2) инкубировали смесь 5 мин при  $65^{\circ}$ С, 3) охлаждали термостат с пробами до  $42^{\circ}$ С, 4) добавляли коммерческий буферный растворный раствор

и обратную транскриптазу согласно указаниям производителя и инкубировали смесь 1 час при 42°С, 5) для остановки реакции и разрушения остаточной РНК прогревали пробы при 75°С 10 мин.

#### 2.6.4 Проведение ПЦР и выделение фрагментов ДНК из геля

Для проведения реакции использовали коммерческие ДНК-зависимые ДНК-полимеразы Taq (Fermentas/Thermo Scientific), Dream-taq (Fermentas/ Thermo Scientific) или Pfu (Fermentas/ Thermo Scientific), dNTP (Fermentas/Thermo Scientific) и соответствующие коммерческие буферный растворные растворы согласно рекомендациям производителя. Обычно проводили 35 циклов со следующим чередованием температур инкубации: 15 сек. 95 °C; 15 сек 55 °C; 30 сек 72 °C. Перед проведением циклов смесь прогревали при 95 °C в течении 3 мин. После проведения циклов смесь инкубировали при 72 °C в течении 10 мин для достройки цепей. Полученные фрагменты ДНК отделяли от праймеров и матрицы или в 1%-ном агарозном геле из агарозы производства Amersham, или в 0,7%-ном геле из легкоплавкой агарозы (Invitrogen). Фрагмент нужной длины вырезали из геля и выделяли с помощью кита QIAquick Gel Extraction Kit (Qiagen), если использовали 1%-ный агарозный гель. Если использовали гель из легкоплавкой агарозы, фрагмент выделяли следующим методом: плавили гель в двукратном объеме воды, затем проводили трехкратную экстракцию равным объемом насыщенного фенола, концентрирование бутанолом и переосаждение этанолом и ацетатом аммония, как описано ранее для плазмидной ДНК.

#### 2.6.5 Кинетика синтеза вирусной РНК

Для определения кинетики синтеза дочерних вирусных РНК культуру клеток *Vero* в шестилуночных или 12-луночных панелях трансфицировали очищенными в градиенте полногеномными транскриптами, несущими исследуемые мутации в домене d. Лунки шестилуночной панели трансфицировали по 200 нг РНК, в 12-луночной панели трансфицировали по 100 нг РНК, в 24-луночной панели по 50 нг НК. Выделяли тотальную РНК из трансфицированных клеток на 0, 10, 12, 16 и 20 часов после трансфекции с помощью Tri-pearentra или реагента Trizol. На каждую временную точку делали три повторности для каждого мутанта. Также выделяли тотальную РНК из пробы без транскрипта. Количество и качество выделенной РНК измеряли спектрофотометрически по соотношению оптической плотности при 260/280 нм. Для определения количества (+) РНК копий проводили обратную транскрипцию и определяли количество копий вирусного генома с помощью ПЦР в реальном времени. Для проведения реакции обратной транскрипции брали 1 мкг выделенной тотальной РНК из каждой пробы. В реакции обратной транскрипции использовали праймер 3ЕР4. В качестве отрицательного контроля брали РНК, выделенную из мок-трансфицированной культуры клеток. Для подготовки калибровочной прямой для ПЦР в реальном времени использовали тотальную РНК культуры клеток, обработанную смесью для трансфекции без РНК, с добавлением серии разведений полногеномного транскрипта дикого типа или с введенной заменой тетрапетли на gCUUGc. В расчетах исходили из следующих данных: молекулярный вес молекулы полногеномного транскрипта (расчет по последовательности) – 241880 Da. При концентрации 243 нг/мкл на 1 мкл исходного раствора приходится  $6.033 \times 10^{10}$  молекул транскрипта. В реакцию обратной транскрипции вносили  $6 \times 10^{10}$ ,  $6 \times 10^{8}$ ,  $6 \times 10^{7}$ ,  $6 \times 10^{6}$ ,  $6 \times 10^{5}$  молекул РНК. Для точности внесения исходный раствор разводили в 4 раза и вносили по 4 мкл на реакцию.

Для определения количества копий вирусной РНК в пробе методом ПЦР в реальном времени использовали набор реагентов Синтол, содержащий термо-активируемую ДНК-зависимую ДНК-полимеразу, раствор нуклеотидов, буферный раствор, раствор MgCl<sub>2</sub>. Для определения количества копий кДНК использовали праймеры PVL1, PVR1 и олигонуклеотидный зонд PVP1 (Таблица 2). Реагенты смешивали согласно инструкции производителя. ПЦР проводили на приборе «7500 Real Time PCR System» Applied Biosystems. Для получения калибровочной прямой в качестве матрицы в ПЦР использовали кДНК, полученную на предыдущем этапе с использованием соответствующих разведений транскрипта.

# 2.7 Получение рекомбинантного вирусного белка 3CD и эксперименты по связыванию oriL *in vitro*

#### 2.7.1 Экспрессия и очистка вирусного белка 3CD

Для очистки рекомбинантного 3CD была использована Ni-хелатная хроматография (см. материалы и методы). Применение опубликованных ранее методов выделения и очистки 3CD (Xiang *et al.*, 1995b; Roehl *et al.*, 1997; Parsley *et al.*, 1999) не позволяло получить препарат с высокой активностью. Это связано с низкой растворимостью 3CD и, возможно, с неспецифическим взаимодействием с PHK E.coli (Рисунок 20 Б). Эти проблемы были решены в ходе совместной работы в лаборатории M.B. Гарбер (ИБ РАН). Для увеличения солюбилизации 3CD в буферном растворе для лизиса была использована высокая концентрация NaCl (1 M). Для того, чтобы избежать соочищения с 3CD, как PHK-связывающим белком, рибонуклеиновой кислоты *E.coli* в буферный

раствор для лизиса добавляли MgCl<sub>2</sub> до концентрации 100 mM, что приводило к разрушению клеточных PHK. В буферных растворах для аффинной хроматографии MgCl<sub>2</sub> не использовался. Такой способ выделения позволил получать препараты 3CD с достаточной концентрацией (до 300 нг/мкл) и стабильной высокой эффективностью взаимодействия с *ori*L (полное связывание *ori*L в молярном соотношении 1:12). В буферном растворе для постановки реакции использовали 250 mM NaCl. При более низких концентрациях соли белок 3CD выпадал в осадок. При попытке получить более концентрированные препараты в наших условиях 3CD также выпадал в осадок (Рисунок 20 Б). Анализ фракций в ДCH-ПААГ из разных этапов очистки и состав фракций элюции представлен на рисунке 20.



Рисунок 20. Выделение и очистка рекомбинантного 3CD. А – этапы выделения 3CD, В – соотношение растворимой и нерастворимой фракций 3CD после диализа при рекомендуемых в литературе условиях (25 mM KCl, Hepes-KOH pH 7.5, 2 mM DTT, 2.5 mM MgCl<sub>2</sub>).

Штамм *E.coli* JM109 трансформировали плазмидой pQE60-3CD и высевали на чашки с 1.2% агаром на среде SOB, содержащей 20 mM глюкозы. Экспрессия с pQE60 подавлена в присутствии глюкозы за счет наличия в плазмиде двух Lac-операторов. На следующий день выросшие колонии смывали в 200 мл SOB, содержащий 20 mM глюкозы и подращивали при покачивании при температуре 37°C до показателей оптической плотности 0,5-0,6 при 590 нм. Суспензию клеток осаждали центрифугированием при 3000 оборотов/мин и 4°C на центрифуге Bekcman, ротор JA7,5. Осадок ресуспендировали в 200 мл SOB без глюкозы с 2 mM индуктора экспрессии IPTG. Суспензию инкубировали ночь при комнатной температуре при покачивании. После инкубации суспензию осаждали центрифугированием, как было описано выше. Осадок ресуспендировали в 10 мл буферного раствора 1 (Tris-HCl 50 mM, pH7.5, NaCl 1 M, MgCl<sub>2</sub> 100 mM, 4 mM меркаптоэтанол,

2 mM PMSF, лизоцим 7 мг/мл). Инкубировали при  $+4^{\circ}$ C 40 мин. Клетки разрушали ультразвуком 3 раза по 15 секунд с промежуточным охлаждением в ледяной бане. Клеточный дебрис осаждали центрифугированием при 15000 оборотов/мин при  $+4^{\circ}$ C на центрифуге Beckman, ротор JA15. Лизат наносили на колонку с носителем His-select 1 мл (Invitrogen). Промывали 15 мл буферного раствора 2 (50 mM Tris-HCl, pH7.5, 500 mM NaCl, 5% глицерин, 4 mM β-меркаптоэтанол. Затем промывали 15 мл того же буферного раствора, содержащего 15 мМ имидазола. Элюцию осуществляли тем же буферным раствором, содержащим 300 мМ имидазола. Для взаимодействия с PHK белок переводили в буферный раствор 3 (20 mM Tris-HCl pH 8, 250 mM NaCl, 5% глицерин, 2 mM DTT) методом гельфильтрации на колонках, содержащих 5 мл Sepahdex G-75. Концентрацию белка оценивали методом Бредфорд (Bradford, 1976), чистоту – с помощью электрофореза в 12.5% ПААГ по Лэмли (Laemmli, 1970).

#### 2.7.2 Получение антител мыши к рекомбинантному белку 3CD

Иммунизацию беспородной белой мыши проводили 25 мкг рекомбинантного белка 3CD. Предварительно белок осаждали насыщенным раствором сульфата аммония и ресуспендировали в PBS. Полученную суспензию диализовали против 200 объемов физиологического раствора. Суспензию смешивали с адьювантом Фройда (Calbiochem) в соотношении 1:1. Смесь взбивали до густой пены и вводили мышам в холку. Следующий этап иммунизации проводили через 10 дней. Последующие 4 этапа иммунизации проводили с интервалом в 7 дней. Шестое введение делали через 5 дней без адъюванта и еще через неделю привили 0,5 мл перевиваемой опухоли карциномы G180. Через 10 дней была взята асцитная жидкость и кровь. Асцитную жидкость и кровь центрифугировали 20 мин при +4°C при 3000 оборотов/мин, ротор JA 7,5. Надосадочную жидкость смешали с равным объемом глицерина и хранили при – 20°C.

Полученные мышиные сыворотки были апробированы в ходе исследования сывороток, принадлежащих людям, заболевшим полиомиелитом в ходе вспышки в Таджикистане в 2010 году, к неструктурным белкам полиовируса (Yakovenko *et al.*, 2014). Во всех исследованных сыворотках были обнаружены антитела к вирусному белку 3CD.

#### 2.7.3 **Метод EMSA**

Эффективность взаимодействия рекомбинантного 3CD и *oriL* оценивали по результатам электрофореза реакционных смесей в 4% ПААГ, содержащем 5% глицерина и 2-кратный Трисборатный буферный раствор (100mM Tris, 100mM борат, pH 8.0). В реакционных смесях молярное соотношение РНК:белок было примерно 1 к 12 (128 нМ *oriL* к 1541 нМ 3CD). Для проведения реакции смешивали 1 мкл РНК (100 нг/мкл), 15 мкл раствора белка (290 нг/мкл), 1,6 мкл MgCl<sub>2</sub> 25 мМ. Реакцию проводили при 32°C, 15 мин. По окончании реакции добавляли 1,7 мкл краски (50% глицерин с 0,001% бромфеноловым синим) и наносили на гель. Предварительно осуществляли префорез геля при 200 V до уменьшения силы тока в два раза. Электрофорез осуществляли с проточным охлаждением. Комплекс визуализировали двумя методами: 1) окрашиванием РНК бромистым этидием 2) с помощью Вестерн-блот с антителами к 3CD. Для окрашивания бромистым этидием гель по окончании электрофореза помещали в 2 кратный буферный раствор ТВЕ, содержащий бромистый этидий (0.2 мкг/мл), инкубировали несколько мин и визуализировали в ультрафиолетовом излучении.

#### 2.7.4 Визуализация комплексов 3CD/oriL методом Western-blot

Для окрашивания с помощью антител сразу по окончании EMSA осуществляли мокрый перенос на нитроцеллюлозную бумагу. Затем мембрану инкубировали в 5%-ном молоке в буферном растворе TBS-T (20 mM Tris-HCl pH7.5, 0,1% Tween20, 150 mM NaCl) при +4°C в течение ночи. На следующий день мембрану инкубировали час на качалке с мышиными антителами к белку 3CD (см. ниже) в буферном TBS-T в разведении 1:1000. После 3 промывок буферный растворным раствором TBS-T (по 5-10 мин на качалке) инкубировали на качалке час с антителами к мышиным иммуноглобулинам (IgG мыши, коньюгат с пероксидазой хрена, Promega) в разведении 1:2500. После трех промывок буферный растворным растворным раствором TBS-T (по 5-10 мин на качалке) визуализировали с помощью субстрата пероксидазы хрена (ECL visualization kit, GE Healthcare фирма GE Healthcare это фирма).

#### 2.8 Использованное программное обеспечение.

В работе были использованы следующее программное обеспечение:

Скрипты, написанные на языке Python и любезно предоставленные А.Н. Лукашевым – для поиска полногеномных энтеровирусных последовательностей и последующей модификации полученной выборки.

NCBI BLAST – для поиска нуклеотидных последовательностей в базе данных NCBI и их выравнивания (Altschul *et al.*, 1997; "Database resources of the National Center for Biotechnology Information.," 2014).

VMD – для визуализации файлов .pdb (из базы данных PDB), содержащих информацию о пространственной структуре молекул РНК и белков (Humphrey, W., Dalke, A. and Schulten, 1996).

OneDScan – для определения интенсивности окрашивания пятен после Вестерн-блот (ScanalyticsInc, Fairfax, USA).

#### 3. Результаты

## 3.1 Отбор жизнеспособных вариантов полиовируса из набора случайных

Для определения множества последовательностей апикального участка домена d репликативного элемента oriL, способных поддерживать эффективную репликацию вируса полиомиелита, был использован следующий подход (Рисунок 17): 1) синтезирован олигонуклеотид SolD (Таблица 2), содержащий 8 случайных нуклеотидов (с 61-го по 68-ой нуклеотид генома вируса полиомиелита) в контексте последовательности генома полиовируса дикого типа; 2) олигонуклеотид SolD был использован для получения ПЦР-фрагмента, который затем был клонирован в плазмиду pT7PV1Rib(+)MS по сайтам рестрикции ApaI и SpII; 3) полученные плазмиды (одиночные плазмидные клоны или в составе пулов) были использованы в реакции транскрипции in vitro для получения РНК-копий генома и трансфекции культуры клеток Vero; 4) если транскрипт давал жизнеспособное вирусное потомство, то после трансфекции образовывались бляшки (образование бляшек наблюдали до восьмого дня после трансфекции). Бляшки отбирали и использовали для следующих этапов исследования. Для полученных вирусов были определены нуклеотидные последовательности на участке генома, соответствующего клеверному листу (с 1 по 110 нуклеотиды полиовирусного генома). Так была получена информация, какие последовательности из набора случайных позволяют вирусу оставаться жизнеспособным. Схема эксперимента представлена на рисунке Рисунок 22.

#### 3.1.1 Оценка качества олигонуклеотида, использованного для рандомизации

Качество рандомизации олигонуклеотида, использованного для получения плазмид с рандомизированным участком, будет отражаться на спектре отбираемых вариантов и будет влиять на конечный результат. Для оценки качества рандомизации мы взяли 41 случайно отобранный плазмидный клон и для каждой определили нуклеотидную последовательность, соответствующую олигонуклеотиду, использованному для рандомизации. Степень рандомизации оказалась удовлетворительной – частота встречаемости различных нуклеотидов в положениях 61-68 генома варьировала от 12% до 41% с медианой 24% (Таблица 3).

Таблица 3. Частоты встречаемости нуклеотидов в рандомизированном участке у 41 случайно отобранных плазмид. Положения с повышенной (более 36%) частотой встречаемости отмечены серым.

	положение в рандомизированном участке плазмиды							
Нуклеотид	1	2	3	4	5	6	7	8
	Спи	раль	Тетрапетля			Спираль		
Α	39	20	20	24	24	29	15	24
Т	22	22	24	32	34	22	15	39
С	24	41	34	20	12	24	32	20
G	15	17	22	24	29	24	39	17

В отсеквенированном участке только у 10 клонов не было обнаружено дополнительных замен, делеций или вставок в не рандомизированной области олигонуклеотида. В остальных клонах обнаружены: замены одного нуклеотида (12 клонов), вставки одного или двух нуклеотидов (18 клонов, 6 из которых содержали вставку G в последовательность GGGG в положениях 14-17 генома, в четырех случаях в паре с другой вставкой), делеции протяженностью от 1-го до 41-го нуклеотида (1 клон с делецией одного нуклеотида, 2 клона с делециями двух разных нуклеотидов, 2 клона с делециями 3-х и 4-х нуклеотидов, и один клон с двумя делециями – одного нуклеотида и продолжительной делецией 41-го нуклеотида). Изменения отмечены на Рисунок 21.



Рисунок 21. Дефекты, обнаруженные в консервативной области олигонуклеотида, использованного для рандомизации. Вставки - отмечены стрелками. Делеции отмечены красным. Точечные мутации отмечены синим. В скобках указаны количество случаев встречи данного изменения у отсеквенированных плазмид.

#### 3.1.2 Отбор жизнеспособных вариантов полиовируса.

Обнаруженные дефекты в консервативной области олигонуклеотида могут снижать жизнеспособность соответствующих вариантов и сказываться на результатах эксперимента. Чтобы оценить, как дефекты в константной части олигонуклеотида могут влиять на жизнеспособность соответствующих вариантов, мы исследовали два вируса с наиболее часто встречающимися модификациями. Первый со вставкой дополнительного С в последовательность в положениях 23-25 генома (апикальная часть домена b, Рисунок 21). Второй со вставкой дополнительного G в положениях 14-17 генома и вставкой С, аналогичной вставке у первого вируса. Первый вирус сохраняет вставку до второго пассажа. Второй сохраняет обе вставки до пятого пассажа. Таким образом, эти наиболее часто встречающиеся вставки не будут влиять на жизнеспособность содержащих их вариантов полиовируса и, соответственно, не будут оказывать существенного влияния на многообразие отбираемых вариантов.

Восемнадцать плазмидных клонов с известной последовательностью на участке, соответствующей нуклеотидам 1-110 полиовирусного генома, были использованы в реакции транскрипции. Полученные неочищенные транскрипты были использованы для трансфекции культуры клеток *Vero*. Только два транскрипта из 18 оказались способными к формированию бляшек после трансфекции (см. Таблица 4). В обоих случаях прилегающие к тетрапетле фланкирующие пары были комплементарны. Остальные шестнадцать клонов имели либо обе пары (девять клонов), либо одну из пар (семь клонов) разрушенными (Таблица 4). Мы можем высказать предположение, что для жизнеспособности полиовируса важно формирование пар нуклеотидов в прилегающих к тетрапетле участках.

Из индивидуальных плазмидных клонов было сформировано 34 пула по 5 вариантов, представляющих собой эквимолярные смеси препаратов случайно выбранных плазмидных клонов. Продукты транскрипции пулов были использованы для трансфекции культуры клеток *Vero*. Жизнеспособные варианты были обнаружены только в четырех пулах (для трех из них определена нуклеотидная последовательность, соответствующая *ori*L). Сравнение нуклеотидной последовательности вирусов из бляшек с последовательностью исходных плазмид пула показало, что жизнеспособны варианты, имеющие не разрушенные фланкирующие пары (таблица 4).

Таблица 4. Последовательности апикального участка домена d генома полиовируса в плазмидных клонах, не давших и давших жизнеспособные варианты при индивидуальной трансфекции и трансфекции пулами по 5 вариантов.

Удельная инфекционность	Последовательность рандомизированного участка плазмиды	Наличие/отсутст вие изменений*	Описание изменений		
	<u>a</u> aCTGCt <u>a</u> **	нет мутаций			
	tgCGAAc <u>c</u>	3 вставки и делеция	вставки G21, A27, C69, A70, делеция G14		
	<u>aa</u> TAGC <u>ga</u>	вставка и делеция	вставка А в последовательность АААА в положениях 3-6 генома, делеция А88		
	<u>ta</u> GGAT <u>gc</u>	вставка и	мутация Т11С, вставка G в положение 21		
	<u>c</u> cGCGTg <u>a</u>	мутация	С40Т мутация		
	<u>cc</u> ACTA <u>ct</u>	вставка	вставка G33		
Нет бляшек (при трансфекции	<u>c</u> tGTACg <u>a</u>	вставка и две делеции	делеции А90 и С100, А4Т мутация,		
не более 0,3 мкг	<u>ta</u> CTTT <u>gt</u>	нет мутаций			
РНК на флакон)	<u>aa</u> CGCA <u>gg</u>	нет мутаций			
1 /	<u>ac</u> GGTC <u>ca</u>	нет мутаций			
	c <u>t</u> TACA <u>c</u> g	вставка	вставка G в последовательность GGG в положениях 14-17 – не летально		
	<u>tc</u> GTAA <u>at</u>	нет мутаций			
	<u>ag</u> CGAG <u>gg</u>	Вставка	вставка G в положение 48		
	<u>ga</u> CTGA <u>aa</u>	нет мутаций			
	<u>a</u> gAAGGc <u>c</u>	нет мутаций			
	<u>c</u> cGAGCt <u>g</u>	Мутация	замена С12Т		
	ccTATGgg	две вставки	вставки G в положения 21 и 41		
	(плазмидный клон cSld15)				
1-2 бляшки,	acCCTTgt	нет мутаций			
(при трансфекции	(плазмидный клон cSld12)	-			
не более 0,3 мкг	atTTACgt (Пул pV5)	нет мутаций			
РНК на флакон)	agCTGAct (Пул pV13-5)	вставка	вставка С в последовательность ССС в положениях 23-25 (домен b)		
	atAGCAat (Пул pV15)	нет мутаций			

\*- изменения в плазмидных клонах на участке с 1 по 110 нуклеотиды генома полиовируса.

\*\* - жирным и подчеркиванием отмечены разрушенные фланкирующие пары.

Использование индивидуальных плазмидных клонов и пулов из 5 вариантов оказалось малоэффективным, так как давало очень низкий выход жизнеспособных вариантов. Для увеличения эффективности были сформированы пулы большего размера. Для этого выросшие после трансформации *E.coli* колонии были случайно объединены в пулы по 20, 100, 300, 900 и 2500

вариантов. Препараты плазмидной ДНК из пулов использовали для транскрипции *in vitro*. Схема эксперимента представлена на Рисунок 22.



Рисунок 22. Схема эксперимента по отбору жизнеспособных вариантов полиовируса из набора случайных вариантов.

В случае пулов, содержащих 20, 100, 300 и 900 вариантов, для трансфекции использовали продукт транскрипции без дополнительной очистки (см. Материалы и методы). Концентрацию РНКтранскрипта в транскрипционной смеси определяли по интенсивности окрашивания бромистым этидием в агарозном геле. В случае пулов по 2500 вариантов РНК-транскрипт перед трансфекцией очищали в сахарозном градиенте. Количество транскрипта после очистки определяли спектрофотометрически.

На основании количеств внесенной РНК и количества образовавшихся бляшек можно оценить удельную инфекционность пулов и количество инфекционных вариантов во всей совокупности вариантов в пуле (Таблица 5). Удельная инфекционность (S) это количество образовавшихся бляшек (обозначим n), нормированное на количество внесенной РНК в мкг (обозначим m), то есть S=n/m. Примем допущение, то все жизнеспособные/инфекционные (?) варианты имеют удельную инфекционность, сравнимую с транскриптом, последовательность которого соответствует последовательности генома дикого типа. Тогда удельную инфекционность пула с одним жизнеспособным/инфекционным вариантом (ожидаемую удельную инфекционность, обозначим T) можно оценить, разделив удельную инфекционность транскрипта, последовательность которого соответствует последовательности генома полиовируса дикого типа (обозначим СБОЕ/мкг), на количество вариантов в пуле (обозначим А). Это допущение с некоторым занижением позволит оценить количество инфекционных вариантов в пуле (N), делением удельной инфекционности пула (S) на его ожидаемую удельную инфекционность (T), то есть N=S/T. Таким образом, количество инфекционных вариантов в пуле будет равно:

 $N=S/T=[n\times A]/[m\times C_{\text{БОЕ/MKT}}]$ 

Для расчета удельной инфекционности и количества инфекционных вариантов выбирали разведения с количеством бляшек от 5 до 50 (выделены жирным шрифтом в таблице 5). Данные в таблице показывают, что ожидаемо все пулы имеют удельную инфекционность значительно ниже, чем полногеномный транскрипт и что инфекционные варианты содержат пулы от 300 вариантов (Таблица 5).

Кол-во вариантов в пуле (А)	Кол-во внесенной РНК на флакон т (мкг)	Кол-во бляшек на флакон на 3-й день п. т.ª	Среднее кол-во бляшек на флакон (n <sup>6</sup> )	Удельная инфекционн ость пула (S, БОЕ/мкг)	Теоретическая удельная инфекцион- ность (Т)	Кол-во инфекцион- ных. вариантов в пуле (N)
20	<b>2</b> 0,4	<b>4,0</b> 0,0	2	1	3575	менее 1
100	2 0,4	0, 1 <b>3, 3</b>	3	7.5	715	менее 1
300	2 0,4 <b>0,08</b>	нс <sup>в</sup> нс 102, 63	82.5	1031.25	238	4.3
900	2 0,016	нс 15, 11	13	812.5	79.44	10,23
2500 (1)	0,25 0,05 <b>0,01</b> 0,002	нс 28, 19 <b>6, 4</b> 0, 0	5	500	28.6	17.49
2500 (2)	0,25 <b>0,05</b> 0,01 0,002	30, 14 11, 2 0, 7 0, 0	6.5	130	28.6	4.55
2500 (3)	0,25 0,05 0,01 0,002	57, 28 14, 10 1, 0 0, 1	12	240	28.6	8.4

Таблица 5 Частота встречаемости инфекционных вариантов в пулах по 2500 вариантов.

а - после трансфекции

б – значения количества бляшек на флакон, взятые для расчета среднего количества бляшек на флакон, выделены жирным в – не сосчитать, количество бляшек во флаконе слишком велико

Более одного инфекционного варианта полиовируса содержат пулы размером от 300 вариантов. В пулах по 20 и 100 вариантов мы наблюдаем бляшки на третий день, однако удельная инфекционность (S) этих пулов более чем на 2-4 порядка ниже, чем могла бы быть, если бы в пуле

был хотя бы один инфекционный вариант (Т) (Таблица 5). Такая низкая удельная инфекционность при наличии бляшек свидетельствует о присутствии в пуле слабо или квазиинфекционных вариантов, которые для восстановления высокой инфекционности нуждаются в дополнительных изменениях в геноме.

## 3.1.3 Разнообразие последовательностей в апикальном участке домена d у экспериментально отобранных вирусов

Всего мы получили 151 жизнеспособный вирус и из них для 62-х определили последовательность генома на участке с 1 по 110 нуклеотиды. Бляшки после трансфекции появлялись на разные дни и имели разный размер. В крупных и ранних бляшках было достаточно материала для определения нуклеотидной последовательности интересующего нас участка генома. Определение последовательности генома вирусов из более мелких бляшек было затруднено из-за недостаточного количества вирусной РНК. Поэтому вирусы из более мелких и поздних бляшек были дополнительно размножены. Для двух вариантов потребовалось 2-3 раунда размножения.

Из 62-х определенных последовательностей 39 были уникальны (Рисунок 23). Оказалось, что любой нуклеотид может занимать любую позицию в исследованном октануклеотиде, кроме U в четвертой позиции тетрапетли. Результаты представлены на рисунке 23 и в таблице 6.

F	РV1 штамм Ма	honey	AU UG рандом	CG GU	]	
r	рандомизиров	анная N		N <sub>5</sub> N <sub>6</sub> N <sub>7</sub> N <sub>8</sub>		
	отобранны варианты	e G U	G A G G U C A C U C	A A U A U CG G	A C G U	
	консенсу	C YNMG				
	спаренные разрушенн фланкирующие фланкирую пары пары		консенсус YNUG	другие тетрапетли	гетерогенные тетрапетли	
	ccUGCGgg guUGCGac uaUUCGua auUACGgu <sup>a,6</sup> uuUACGaa ccUACGgg aaCCCGuu ucCCCGgg acUCCGgu <sup>a</sup> cuUCAGgg uuUNCGaa <sup>8</sup> ccCGCGgg auUUAGgu	gcUACGgg a <u>u</u> UGCG <u>u</u> u <u>a</u> cCCCGgg <u>gc</u> CCCG <u>a</u> c <u>gc</u> CCAG <u>a</u> c	gcCGUGgu uuUAUGga <sup>a</sup> gcUAUGgc ccUAUGgg <sup>a,r</sup> acCCUGgu uaUCUGua cuCAUGgg agUUUGcu acUUUGgu	GSYA auGCUAgu uuGCUAaa aaGGCAuu auGGCAau Het Kohcehcyca auAGCAau agCUGAcu <sup>6</sup> gcUUGGgc guUUGGgc uaGCACua	aaUUNGuu aaCA(G/C)Guu uuUA(C/U)Gag	

Рисунок 23. Нуклеотидная последовательность апикального участка домена d у экспериментально отобранных вирусов. <sup>а</sup> – вставка G в последовательность GGGG в положениях 14-17 генома, <sup>6</sup>– вставка C в последовательность ССС в положениях 23-25 генома, <sup>в</sup>–замена A26G, <sup>г</sup> – вставка G в положении 40 генома.

Таблица 6. Частоты встречаемости различных нуклеотидов в положениях рандомизированного участка среди 39 уникальных последовательностей (приведен процент встречаемости). Нуклеотиды с повышенной частотой встречаемости, относительно средней, в положениях с 61 по 68 генома отмечены зеленым, с пониженной частотой встречаемости отмечены красным.

	положение в рандомизированном участке вируса								
нуклеотид	1	2	3	4	5	6	7	8	
	Спи	раль	Тетрапетля			Спираль			
Α	44	14	3	25	11	17	17	17	
U	19	39	56	19	31	0	17	42	
С	17	42	28	33	50	3	5	14	
G	19	6	14	22	8	81	61	28	

Повышенная частота встречаемости нуклеотидных остатков аденина в первом положении, цитозина во втором, гуанина в седьмом и урацила в восьмом в экспериментально отобранных последовательностях коррелирует с частотой встречаемости соответствующих нуклеотидных остатков в этих положениях в плазмидных клонах (Таблицы 3 и 6). Поэтому наблюдаемые частоты встречаемости в перечисленных положениях октануклеотида, скорее всего, обусловлены распределением частоты встречаемости нуклеотидных остатков в плазмидах, а не влиянием естественного отбора (Таблицы 3 и 6). Однако легко заметить, что в вирусных последовательностях частота встречаемости нуклеотидов в ряде случаев отличается от распределения частот встречаемости нуклеотидов в плазмидных последовательностях. Это может свидетельствовать о наличии положительного и отрицательного отбора в ходе эксперимента. Анализ показывает селекцию U в третьем положении рандомизированного октонуклеотида (первом положении тетрапетли) в сочетании с G в шестом положении октонуклеотида (последнем положении тетрапетли) и селекцию С в пятом положении октонуклеотида (третьем положении тетрапетли). Во втором положении тетрапетли (четвертом положении октануклеотида) не обнаруживается никаких предпочтений. Наблюдается предпочтение неоднозначной пары u-g (в англоязычной литературе wobble pair – специфические, не Уотсон-Криковские пары) в качестве фланкирующей к тетрапетле.

Большинство отобранных вариантов можно разделить на три группы, тетрапетли в которых принадлежат соответственно трем консенсусам YNMG, YNUG и GSYA (Рисунок 23).

Восемнадцать из 39 уникальных последовательностей объединяются консенсусом YNMG (Y=U/C, N=любой, M=A/C). Тетрапетли этого консенсуса широко распространены в рибосомальных PHK (Woese *et al.*, 1990). Все изученные тетрапетли этого консенсуса формируют пространственную структуру класса UNCG (см. Обзор литературы, раздел 1.3). Некоторые тетрапетли, принадлежащие консенсусу YNMG, встречались среди жизнеспособных вариантов несколько раз с разными вариантами фланкирующих пар, в том числе с разрушенными фланкирующими парами: UACG и СССС с четырмя вариантами фланкирующих пар, UGCG с тремя вариантами.

Вторая группа из 9 последовательностей формировала консенсус YNUG (Рисунок 23). К этому консенсусу могут относиться представители двух разных структурных пространственных классов тетрапетель (UNCG и gCUUGc, см. обзор литературы раздел 1.3). В нашей выборке были обнаружены представитель класса UNCG, тетрапетля UUUG. Тетрапетля UUUG с фланкирующими парами ас...gu формирует пространственную структуру класса UNCG (Cheong, Varani & Tinoco, 1990; Varani *et al.*, 1991).

Третья группа из четырех последовательностей - auGGCAau, auGCUAgu,

uuGCUAaa - составляла консенсус GSYA (S=G/C). Для тетрапетли GCUA с фланкирующими парами au...gu методом ЯМР была продемонстрирована пространственная структура класса UNCG (Melchers *et al.*, 2006).

Остальные восемь тетрапетель включают ucAGCAga, agCUGAcu, gcUUGGgc, guUUGGgc, uaGCACua, uuUA(C/U)Gag, aaCA(G/C)Guu, aaUUNGuu. Эти тетрапетли не объединяются в какойлибо осмысленный консенсус и, кроме тетрапетли AGCA, не имеют аналогов в литературе.

Тетрапетля AGCA с фланкирующими парами uc...ga узнается РНКазой III дрожжей *Saccaromyces cerevisiae* практически с той же эффективностью, как и тетрапетли пространственного класса AGNN (Butcher *et al.*, 1997; Lebars *et al.*, 2001; Wu *et al.*, 2001, 2004). На основании этого можно предположить, что, auAGCAgu может формировать пространственную структуру класса AGNN. Пространственная структура этого класса отличается от таковой класса UNCG (Lebars *et al.*, 2001).

Таким образом, три многочисленные группы последовательностей содержат тетрапетли, для которых показана пространственная структура класса UNCG. При этом представители распространенных пространственных структурных классов GNRA и gCUUGc обнаружены не были.

## 3.1.4 Разнообразие последовательностей тетрапетли домена d и фланкирующей пары у вирусов вида Enterovirus C

В базе данных pubmed.com на момент данного исследования обнаружено всего 747 полногеномных последовательности, принадлежащих виду Enterovirus C, содержащих 13 уникальных тетрапетель и две тринуклеотидных петли в апикальном участке домена d *ori*L (Таблица 7). Из них 9 принадлежит консенсусу YNMG и 4 консенсусу YNUG (Таблица 7). Многочисленные близкородственные геномы, принадлежащие, например, вирусам, выделенным в ходе одной вспышки, будут влиять на оценку распространенности разных последовательности тетрапетель. Чтобы снизить это искажение, из выборки были исключены последовательности, отличающейся от других на 1% и менее. Фильтрация близкородственных вариантов сокращает выборку более чем в два раза, в основном, за счет последовательностей многочисленных изолятов полиовируса из Нигерии (Famulare *et al.*, 2016; Montmayeur *et al.*, 2017). При этом фильтрация близкородственных вариантов практически не влияет на распределение последовательностей тетрапетель по распространенности и на их разнообразие по фланкирующим парам (Таблица 7). Однако в результате фильтрации теряется три уникальных тетрапетли СССС, САUG и UGUG, что указывает
на то, что даже в популяции близкородственных вирусов, например, в ходе одной вспышки, апикальная тетрапетля домена d может меняться.

Наиболее распространенными вне зависимости от того, проведена или не фильтрация по близкому родству, были тетрапетли (по убыванию встречаемости) САСС, UGCG, UACG (Таблица 7). Эти же варианты тетрапетель имеют наибольшее разнообразие по последовательности фланкирующих пар: 5 вариантов у САСС (3 и...g, 1 и...а, 1 с...g), 4 варианта у UGCG и 2 варианта у UACG (Таблица 7).

Таблица 7. Анализ распространенности и разнообразия фланкирующих пар тетрапетель апикального участка домена d oriL у энтеровирусов С.

N	Тетрапетля	Распространенность	Распространенность (фильтр 1%)/принадлежащие полиовирусу	Разнообразие фланкирующих пар	Разнообразие фланкирующих пар (фильтр 1%)			
			YNMG					
1	CACG	254	101/54	5	4			
2	UGCG	232	43/31	4	4			
3	UACG	226	106/64	2	2			
4	CGCG	21	13/6	1	1			
5	CUCG	2	2	1	1			
6	CGAG	2	2	1	1			
7	UCCG	1	1/1	1	1			
8	CCCG	1	0	1	0			
9	CAAG	1	1	1	1			
	YNUG							
10	CGUG	2	2	1	1			
11	CAUG	1	0	1	0			
12	CCUG	1	1	1	1			
13	UGUG	1	0	1	0			
Тринуклеотидные петли								
14 CCG		1	1	1	1			
15	CAG	1	1	1	1			
Всего		747	274	23	19			

Очевидно, что широкая представленность в базе данных последовательностей, принадлежащих полиовирусу (в основном, дериватами вакцинных штаммов) оказывает влияние на результат (Таблица 7). Даже после фильтрации близкородственных вариантов, тетрапетли, принадлежащие к полиовирусам, дают значимый вклад во встречаемость тетрапетель (Таблица 7). Так, тетрапетли UGCG в значительной части (27 последовательностей из 43) принадлежат полиовирусу первого типа (Таблица 7, Таблица 8). При этом 25 из них являются дериватами вакцинного штамма. Тетрапетли UACG почти наполовину представлены полиовирусом второго типа (46 последовательностей из 106, одна последовательность из которых принадлежит вакцинному штамму, а остальные его дериватам) (Таблица 7, Таблица 8). Тетрапетли CACG также почти на половину принадлежат полиовирусам (Таблица 7, Таблица 8).

Таблица 8. Распространенность последовательности апикального участка домена d у полиовирусов после удаления близкородственных штаммов.

Последовательность нонануклеотида	PV1	PV2	PV3	Всего				
	UGCG							
uauUGCGgua	25	3	1					
uacUGCGgua	2	0	0	31				
Всего	27	3	1					
	UACG							
uauUACGgua	10	46	1					
uacUACGgua	7	0	0	64				
Всего	17	46	1					
CACG								
uauCACGgua	9	22	20					
uacCACGgua	2	0	0	E A				
uguCACGgua	0	0	1	54				
Всего	11	22	21	1				
другие тетрапетли								
uauUCCGgua	0	1	0	1				
uauCGCGgua	4	2	0	6				
ВСЕГО	59	74	23	156				

Однако, хотя полиовирусы являются значимой частью выборки, их исключение оставляет в лидерах по распространенности тетрапетли CACG и UACG, так как, как оказалось, эти тетрапетли широко распространены среди вирусов коксаки A (CVA) и других энтеровирусов C (Таблица 7). В свете того, что энтеровирусы вида C свободно обмениваются 5'частями генома, это не является удивительным (Muslin *et al.*, 2015; Bessaud *et al.*, 2016).

Для полиовируса нет строго консервативной последовательности апикального участка домена d: были обнаружены полиовирусы с различными последовательностями – UACG, UGCG, CACG, UCCG и CGCG (Таблица 8), практически независимо от серотипа. Тетрапетли в апикальном участке домена d oriL полиовирусов в базе данных pubmed.com можно описать консенсусом YVCG, где Y=C/U, V=все, кроме U.

Таким образом, спектр вариантов апикальной тетрапетли домена d, полученный в результате отбора жизнеспособных вариантов из случайных, оказался значительно шире, чем таковой у

природных вирусов вида Enterovirus C (Таблица 9). С другой стороны, у природных изолятов были обнаружены тетрапетли, не найденные у вирусов, отобранных экспериментально. Это были широко распространенная у энтеровирусов тетрапетля CACG, тетрапетли CUCG, UGUG и тринуклеотидные петли CCG и CAG (Таблица 9). При этом ни в той ни в другой выборке не были обнаружены последовательности, содержащие в апикальном участке домена d тетрапетли пространственных классов GNRA и gCUUGc.

Таблица 9 Сравнение спектра последовательностей, принадлежащих природным энтеровирусам вида Enterovirus C и вирусов, отобранных из набора рандомизированных по апикальному участку домена d oriL.

Консенсус	YNCG							
Тетрапетля	UACG	UGCG	UUCG	UCCG	CACG	CGCG	CUCG	CCCG
SELEX	3	3	1	1	0	1	0	4
EV-C	106	43	0	1	101	13	2	<u>1</u>
Консенсус				YI	NAG			
Тетрапетля	UAAG	UGAG	UUAG	UCAG	CAAG	CGAG	CUAG	CCAG
SELEX	0	0	1	1	1	1	0	1
EV-C	0	0	0	0	1	2	0	0
Консенсус				YI	NUG			
Тетрапетля	UAUG	UGUG	UUUG	UCUG	CAUG	CGUG	CUUG	CCUG
SELEX	3	0	2	1	1	1	0	1
EV-C	0	<u>1</u>	0	0	<u>1</u>	2	0	1
Консенсус				GS	GSYA			
Тетрапетля	GCUA	GGCA						
SELEX	2	2						
EV-C	0	0						
Консенсус	Нет консенсуса							
Тетрапетля	UUGG	AGCA	CUGA	UUGG	GCAC	CCG	CAG	
SELEX	2	1	1	1	1	0	0	
EV-C	0	0	0	0	0	1	1	

Аминокислоты в положениях 154-156 белка 3С значительно более консервативны - почти во всех изолятах после фильтрации близкородственных вариантов находится аминокислотная последовательность TGK (Таблица 10). Таблица 10 Анализ распространенности и разнообразия аминокислотной последовательности 153-156 белка 3С у энтеровирусов С.

тетрапетля	Распространенность	Аминокислотный триплет в положениях 154-156					
	(фильтр 1%)	TGK		IGK		PGK	
UGCG	43	CTGK	38			APGK	1
		ATGK	3				
		STGK	1				
CACG	101	CTGK	94				
		ATGK	6				
		STGK	1				
UACG	106	CTGK	101	CIGK	1		
		ATGK	4				
UCCG	1	CTGK	1				
CGCG	13	CTGK	11				
		STGK	2				
CUCG	2	CTGK	1				
		ATGK	1				
CGAG	2	CTGK	1				
		ATGK	1				
CGUG	2	STGK	2				
CCUG	1	STGK	1				
CAAG	1	ATGK	1				
CAG	1	CTGK	1				
CCG	1	CTGK	1				
Всего	274	272			1		1

На основании этих данных, на основании присутствия в трех группах экспериментально отобранных вирусов тетрапетель, принадлежащих пространственному классу UNCG и на основании данных литературы о пространственной структуре апикального участка домена d ряда энтеровирусов мы предполагаем, что пространственная структура апикального участка домена d oriL важнее для связывания с вирусным белком 3CD, чем его последовательность. Следствиями из этой гипотезы будут являться 1) разные последовательности в апикальном участке домена d, формирующие структуру подходящего класса, будут с одинаковой эффективностью узнаваться вирусным белком 3CD, 2) последовательности в апикальном 3CD, вне зависимости от последовательности.

## 3.2 Конструирование полиовирусов, содержащих мутации в апикальном участке домена d, и эффективность их репродукции

Связывание белка 3CD и репликативного элемента oriL необходимо для инициации вирусной репликации и, значит, нарушение этого процесса будет негативно сказываться на вирусной репродукции и репликации. Таким образом, следствия из гипотезы о значении пространственной структуры апикального участка домена d для узнавания вирусным белком 3CD могут быть проверены как в серии прямых экспериментов по связыванию oriL с различными последовательностями в домене d вирусным белком 3CD, так и в серии косвенных экспериментов с вирусными геномами, содержащими различные последовательности в апикальном участке домена d.

Мы сконструировали набор вирусных геномов с различными последовательностями в апикальном участке домена d. Кроме того, эксперименты с конструированными геномами позволяют нам убедиться, что обнаруживаемые эффекты обусловлены именно внесенной мутацией, что для отобранных вирусов требует дополнительной проверки. Набор плазмид, содержащих полные геномы полиовируса с заданными мутациями в апикальном участке домена d, был создан на базе плазмиды pT7PV1(Rib+)MS.

Плазмиды с внесенными мутациями были использованы для получения полногеномных транскриптов методом транскрипции *in vitro*. Транскрипты были очищены от плазмидной ДНК и примесей РНК длиннее или короче полногеномного транскрипта с помощью центрифугирования в градиенте сахарозы (см. раздел Материалы и методы). Реконструированные таким образом РНКгеномы были использованы для трансфекции культуры клеток *Vero*. Мы оценивали эффективность репродукции, сравнивая размер бляшек под покрытием после трансфекции культуры клеток *Vero* контрольным транскриптом и транскриптом с исследуемой мутацией. Размер бляшек является интегральным показателем эффективности цикла репродукции нескольких поколений вируса от способности инфицировать клетки до сборки и выхода вирусных частиц и, таким образом, позволяет давать первичную характеристику мутантных вирусных вариантов. В качестве контрольного транскрипта был использован транскрипт, соответствующий полногеномной последовательности вируса полиомиелита штамма Mahoney (далее контрольный транскрипт).

Бляшечный фенотип после трансфекции полученными полногеномными транскриптами представлен на рисунке 24.



Рисунок 24 Бляшечный фенотип после трансфекции полногеномными транскриптами, содержащие различные замены в апикальном участке домена d oriL. Указан день после трансфекции.

Появление бляшек, более мелких, чем у контрольного транскрипта, расценивалось нами, как указание на сниженную жизнеспособность соответствующего транскрипта и, соответственно на негативное влияние внесенных мутаций на жизнеспособность. Пассирование вирусов из мелких бляшек позволяет осуществиться отбору благоприятных мутаций, и, таким образом, дает дополнительную информацию о роли внесенных мутаций или механизмах их компенсации. Генетическая стабильность отобранных из бляшек вирусов приведена в таблице 11.

Последовательность Относительная Изменения после пассажей Размер бляшек<sup>а</sup> удельная Ν Структура тетрапетли и в домене d в 3С фланкирующих пар инфекционность6 Консенсус YNMG, спаренные фланкирующие пары 1 auUGCGgu (w.t.) UNCG крупные 1 нет 2 agUGCGcu крупные нет данных нет данных неизвестно нет данных 3 uuUACGaa нет данных неизвестно крупные 4 auUCAGgu 0.2 нет данных неизвестно крупные auUGAGgu 5 крупные 1 нет данных неизвестно Консенсус YNMG, разрушенная фланкирующая пара 6 auUGCGuu нет данных aaUGCGuu<sup>®</sup> нет неизвестно крупные 7 acCCCGgg маленькие нет данных acCCCGgu нет неизвестно 8 gcCCAGac средние нет данных gcCCAGgc нет неизвестно Консенсус YNUG 9 auUGUGgu 1 нет данных неизвестно крупные 10 acUUUGgu 0.3 YNMG нет нет данных крупные 11 agUUUGcu 0.3 нет данных неизвестно средние нет 12 auCUUGgu 0.8 нет неизвестно крупные нет ag<u>U</u>UUGcu 13 agCUUGcu гетерогенный 0.2 или нет gCUUGc agCU<u>C</u>Gcu Консенсус GSYA 14 T154I auGGCAau точечные 0.3 нет неизвестно 15 T154I aaGGCAuu маленькие <u>нет дан</u>ных нет неизвестно 16 auGGCAgu 0.5 маленькие нет нет данных неизвестно 17 auGGUAgu 0.7 неизвестно средние нет нет данных 18 YNMG auGCUAgu 0.2 крупные нет нет T154I или 19 uuGCUAaa точечные 1.1 неизвестно нет K156R 20 auGCUAau 0.1 нет данных неизвестно крупные нет Структурный класс GNRA и его ревертатнт точечные 21 auGAGAgu 0.005 auGA<u>U</u>Agu GNRA нет на 4-6 день п.т. 22 auGAUAgu 0.7 неизвестно средние нет нет Другие тетрапетли 23 agCUGAcu agCU**CG**cu или T154I гетерогенные не определяли неизвестно 24 auUUA**G**gu auUUACgu гетерогенные не определяли нет данных неизвестно 25 gcUUGGgc gcUU<u>U</u>Ggc маленькие не определяли нет данных неизвестно 26 auGCACau T154I маленькие не определяли нет неизвестно Пентапетля UCAGU T154I, UNCG-подобная точечные P89S +T154I, с выпетленным 27 auUCAG<u>U</u>gu 0.04 нет P89T, P2S, на 4-6 день п.т. основанием 5го нуклеотида N34H

Таблица 11 Бляшечный фенотип, удельная инфекционность и генетическая стабильность сконструированных геномов

а - размер бляшек на третий день после трансфекции, если не указано иное

б - удельная инфекционность, нормализованная на удельную инфекционность полногеномного транскрипта полиовируса штамма Mahoney

в- измененные с пассажами нуклеотиды выделены жирным шрифтом

# Характеристика вариантов с тетрапетлями, принадлежащими консенсусу YNMG и пространственному структурному классу UNCG

Мы сконструировали геномы, содержащие в апикальном участке домена d тетрапетли с разной последовательностью консенсуса YNMG (как обнаруженные так и не обнаруженные в SELEX), для которых показана способность формировать пространственную структуру класса UNCG – UACG, UGCG, UCAG, CCCG (Abdelkafi *et al.*, 1998; Theimer *et al.*, 2003; Zhou *et al.*, 2005; Miller *et al.*, 2014), а также другие тетрапетли, принадлежащие консенсусу YNMG – CCAG, UGAG. Все исследованные транскрипты, в которых фланкирующие тетрапетлю нуклеотиды были способны формировать пары, после трансфекции давали бляшки, сравнимы с бляшками контрольного транскрипта. Транскрипты, в которых одна фланкирующая пара при апикальной тетрапетле домена d была разрушена, после трансфекции давали мелкие бляшки. Этот эффект был менее выражен для транскрипта, содержащего тетрапетлю дикого типа UGCG. Вирусы, полученные из этих бляшек, при пятикратном пассировании приобретали замены, которые приводили к восстановлению фланкирующих пар.

Таким образом, исследованные нами варианты тетрапетель, принадлежащие консенсусу YNMG, способны эффективно поддерживать репродукцию вируса полиомиелита в случае, если фланкирующие пары при тетрапетле не разрушены. Разрушение фланкирующих пар при апикальной тетрапетле домена d оказывают негативный эффект на жизнеспособность полиовируса, однако в случае тетрапетли вируса дикого типа UGCG негативный эффект менее выражен, чем для других исследованных вариантов.

Из восемнадцати экспериментально отобранных вариантов, принадлежащих консенсусу YNMG, пятнадцать содержали нуклеотидный остаток цитидина в третьем положении тетрапетли и только четыре нуклеотидный остаток аденозина (Рисунок 23). Это могло указывать на то, что нуклеотидный остаток аденозина в третьем положении неблагоприятен для жизнеспособности вируса и подвергается отрицательному отбору. Описанный выше реконструированный вариант с тетрапетлей UCAG имеет крупнобляшечный фенотип, что указывает на то, что нуклеотидный остаток аденозина в третьем положении тетрапетли, в случае ее принадлежности консенсусу YNMG, не оказывает негативного влияния на репродукцию вируса. Чтобы оценить влияние нуклеотидного остатка аденозина в окружении последовательности генома вируса дикого типа, мы сконструировали вариант, содержащий тетрапетлю UGAG. Такой вариант имел фенотип не отличимый от таковых для контрольного транскрипта (Рисунок 24, Таблица 11). Таким образом, нуклеотидный остаток аденозина в третьем положении тетрапетли не оказывает негативного влияния на жизнеспособность вируса.

#### Характеристика вариантов с тетрапетлями, принадлежащими консенсусу YNUG

Мы сконструировали геномы, содержащие в апикальном участке домена d тетрапетли с разной последовательностью консенсуса YNUG: обнаруженную в SELEX тетрапетлю UUUG с разной фланкирующей парой и не обнаруженную в SELEX тетрапетлю UGUG. Для тетрапетли UUUG с фланкирующими парами ас-gu показана пространственная структура класса UNCG (Varani *et al.*, 1991).

Все сконструированные транскрипты формировали крупные бляшки на третий день после трансфекции и имели удельную инфекционность, близкую к таковой для контрольного транскрипта (Рисунок 24, Таблица 11). Вирусы, полученные из бляшек после трансфекции, были пятикратно пассированы на культуре клеток *Vero*. Нуклеотидная последовательность участка генома с 1 по 110 нуклеотид у пассированных вирусов соответствовала последовательности транскрипта, использованного для трансфекции (Рисунок 24, Таблица 11). Таким образом, исследованные тетрапетли эффективно поддерживают вирусную репродукцию.

#### Характеристика вариантов с тетрапетлями, принадлежащими консенсусу GSYA

Мы сконструировали геномы, содержащие в апикальном участке домена d обнаруженные в SELEX и созданные заново тетрапетли с разной последовательностью консенсуса GSYA: тетрапетли GCUA и GGCA с разными фланкирующими парами. Для тетрапетли GCUA с фланкирующими парами аu-gu продемонстрирована пространственная структура класса UNCG (Melchers *et al.*, 2006)

Варианты, содержащий тетрапетлю GCUA и фланкирующие пары 5'au-gu3' или 5'au-au3', обладали инфекционностью и бляшечным фенотипом, близкими к таковым для контрольного транскрипта (Рисунок 24, Таблица 11).

Вариант с тетрапетлей GCUA и фланкирующими парами 5'ии-аа3' имел мелкобляшечный фенотип на третий день после трансфекции (Рисунок 24, Таблица 11), Рисунок 27). Секвенирование участков генома вирусов из бляшек на участках, соответствующих области *oriL* (область генома с 1 по 110 нуклеотид), и области, кодирующей 3С (область генома с 5600 по 6300 нуклеотид), выявило аминокислотные замены лизина в положении 156 на аргинин или треонина в положении 154 на изолейцин в белке 3С (Рисунок 24, Таблица 11). Точно такие же аминокислотные замены были описаны в литературе, как компенсаторные, для варианта со значительным нарушением структуры апикального участка домена d (вставка 4 нуклеотидов после 64-го нуклеотида генома) (Andino *et al.*,

1990b). Таким образом, фланкирующие пары 5'uu-aa3' при тетрапетле GCUA оказывают негативное влияние на жизнеспособность вируса, и для восстановления эффективности репродукции необходимы компенсаторные мутации в 3С. При этом фланкирующие пары 5'uu-aa3' сами по себе не оказывают негативного влияния на жизнеспособность вируса (Рисунок 24, Таблица 11). Значит негативный эффект обуславливается сочетанием именно этих фланкирующих пар и тетрапетли GCUA.

Были сконструированы вирусные геномы, содержащие тетрапетлю GGCA, обнаруженную в SELEX, в окружении разных фланкирующих пар. Вне зависимости от последовательности фланкирующих пар транскрипты с тетрапетлей GGCA в апикальном участке домена d давали мелкие бляшки после трансфекции (Puc). При пассировании, вирусы, отобранный из бляшек, приобретали аминокислотную замену T154I в белке 3C, которая известна, как компенсаторная. При этом, искусственная замена третьего C тетрапетли на U возвращает крупнобляшечный фенотип (Puc).

Таким образом, из отобранных тетрапетель, формирующих консенсус GSYA, только GCUA со спиралью, соответствующей последовательности генома вируса полиомиелита дикого типа, способна поддерживать эффективную репродукцию вируса полиомиелита.

# Характеристика отобранных в SELEX вариантов, с тетрапетлями, не принадлежащими консенсусам YNMG, YNUG и GSYA

Мы реконструировали отобранные в SELEX последовательности, не отнесенные к какомулибо консенсусу или структурному классу – gcUUGGgc, auGCACau, agCUGAcu и auAGCAau.

Все реконструированные варианты продемонстрировали мелкобляшечный или гетерогенный с преобладанием мелких бляшек фенотип (Рисунок 24, Таблица 11), фенотип для auGCACau не показан (изображение утеряно). С пассажами эти вирусы приобретали замены, приводящие тетрапетлю к консенсусу YNMG или GSYA, и/или приобретали аминокислотную замену Thr154lle в 3С (Рисунок 24, Таблица 11). Вирус, отобранный из бляшек, после трансфекцией транскриптом, содержащим в апикальном участке домена d тетрапетлю AGCA, эволюционировал двумя путями: 1) либо сначала приобретая замену в тетрапетле, затем в белке 3С и 2) сначала приобретая замену в белке и затем в тетрапетле (Рисунок 25). Таким образом, тетрапетля GGCA оказывается более предпочтительной, чем тетрапетля AGCA, но все же, не способной поддерживать эффективную репродукцию вируса.



Рисунок 25. Два пути эволюции вируса с последовательностью тетрапетли auAGCAau.

Таким образом, последовательности в апикальном участке домена d, не принадлежащие консенсусам YNMG, YNUG или GSUA, не поддерживают эффективную репродукцию вируса полиомиелита.

# Характеристика варианта, содержащего тетрапетлю пространственного структурного класса UNCG и пятый выпетленный нуклеотид

В литературе описана пентапетля, четыре первых нуклеотида которой формируют пространственную структуру класса UNCG, а пятый нуклеотид выпетлен и создает стерические затруднения для взаимодействия с тетрапетлей с 3'-стороны большого желобка (Theimer *et al.*, 2003) (Рисунок 26). Сравнение структур ucUCAGUga (PDB ID: 1Q75) и auGCUAgu (PDB ID: 2EVY), извлеченных из базы данных Protein Data Bank, демонстрирует сходство пространственной структуры тетрапетель GCUA и UCAG: второй нуклеотид тетрапетли находится в малом желобке, третий нуклеотид находится в большом желобке. Отличие – выпетленный в большой желобок пятый нуклеотид пентапетли UCAGU (Рисунок 26). Такая структура представляет интерес в свете существующих гипотез о взаимной ориентации 3CD и домена d *oriL* при взаимодействии.



Рисунок 26. Сравнение ориентации оснований в пентапетле UCAGU (pdb id: 1Q75) и тетрапетле GCUA (pdbid: 2EVY), визуализация с помощью программы VMD (University of Illinois): толстой линией отмечен ход сахарофосфатного остова и второй и третий нуклеотиды у auGCUAgu, и второй, третий и пятый нуклеотиды у ucUCAG(U)ga, тонкими линиями – остальные нуклеотиды. Обе структуры ориентированы 3' стороной тетрапетли к наблюдателю таким образом, что слева малый желобок тетрапетли, справа большой желобок тетрапетли.

Для близких родственников полиовируса, коксакивируса ВЗ и бычьего энтеровируса 1, определены нуклеотиды домена d, вовлеченные во взаимодействие с протеазой 3С (подробнее см. обзор литературы) (Ohlenschläger *et al.*, 2004; Ihle *et al.*, 2005). На основании взаимной ориентации нуклеотидов домена d коксакивируса ВЗ, вовлеченных во взаимодействие с белком 3С, авторы делают предположение, что узнавание осуществляется со стороны малого желобка тетрапетли (Ohlenschläger *et al.*, 2004). Согласно этому предположению, присутствие выпетленного нуклеотида со стороны большого желобка не будет оказывать значительного негативного эффекта на репродукцию полиовируса.

Чтобы проверить это предположение, мы сконструировали вариант, содержащий пентапетлю auUCAGUgu в апикальном участке домена d. В качестве контроля для пентапетли UCAGU мы сконструировали четырехнуклеотидный аналог UCAG без пятого выпетленного нуклеотида. Как и следовало ожидать, фенотип генома с четырехнуклеотидным аналогом UCAG не отличим от такового для контрольного транскрипта (Рисунок 24, Таблица 11). При этом вариант с пентапетлей имел значительно сниженную удельную инфекционность по сравнению с инфекционностью контрольного транскрипта и продуцировал бляшки лишь на седьмой день (Рисунок 24, Таблица 11). Таким образом, пентапетля UCAGU в апикальном участке домена d *oriL* не способна поддерживать эффективную репродукцию полиовируса, что указывает на, что выпетленный пятый нуклеотид со

стороны большого желобка тетрапетли мешает узнаванию 3CD и oriL. Таким образом, узнавание домена d oriL хотя бы частично осуществляется со стороны большого желобка. Для того чтобы исследовать, как будет эволюционировать вариант, содержащий пентапетлю, мы отобрали 7 бляшечных клонов и размножили их. Фенотип полученных вирусов был гетерогенный (не показано). На этом этапе секвенирование участков генома, соответствующих oriL и последовательности, кодирующей 3CD, не выявило дополнительных замен. Мы снова клонировали и размножили все отобранные вирусы. Далее был проведен пассаж с низкой множественностью заражения (0,1 БОЕ/клетку). Фенотип полученных вирусов был гомогенный крупно- или среднебляшечный (не показано). Мы определили последовательности участков генома, соответствующих oriL и 3CD. Все варианты сохранили пентапетлю в апикальном участке домена d и приобрели мутации в 3С. Пять вариантов приобрели описанную Andino et al. компенсаторную замену Thr<sub>154</sub>Ile в белке 3С. Один вирус приобрел дополнительно к замене Thr<sub>154</sub>Ile замену Pro<sub>89</sub>Ser. Другие два вируса не приобрели замену Thr $_{154}$ Ile, но приобрели другие замены: Asn $_{33}$ His или Pro $_{2}$ Ser. Все эти аминокислотные замены находятся непосредственно в районе РНК-связывающей поверхности белка 3С (см. Обзор литературы). Для исследования роли описанных аминокислотных замен необходимо провести дополнительные эксперименты.

Таким образом, вариант полиовируса, содержащий пентапетлю в апикальном участке домена d, не способен к эффективной репродукции и с пассажами приобретает мутации, по-видимому компенсаторные, в РНК-узнающем сайте 3С. На основании этих данных мы делаем вывод, что выпетливание пятого нуклеотида в большой желобок, скорее всего, создает стерические затруднения для эффективного взаимного узнавания *oriL* и 3CD. Мы предполагаем компенсаторную природу обнаруженных замен, однако доказательство этого требует дополнительных экспериментов.

## Характеристика варианта с тетрапетлей пространственного структурного класса GNRA и его ревертанта

В качестве представителя структурного класса GNRA была использована тетрапетля GAGA с фланкирующими парами 5'au-gu3', соответствующими таковым для полиовируса дикого типа. Полученный транскрипт был использован для трансфекции культуры клеток *Vero*. Бляшки появлялись на поздних сроках (6-8 день) (Рисунок 24, Таблица 11).

Удельная инфекционность варианта, содержащего тетрапетлю GAGA, была на 3 порядка ниже, чем удельная инфекционность контрольного транскрипта (50 БОЕ/мкг по сравнению с 3.6×10<sup>4</sup>

#### БОЕ/мкг) (Рисунок 24, Таблица 11).

Таким образом, тетрапетля, принадлежащая пространственному классу GNRA, не способна поддерживать эффективную репродукцию вируса.

Так как вариант, содержащий тетрапетлю GAGA, был слабожизнеспопсобным и бляшки после трансфекции таким вариантом появлялись на поздних сроках, мы ожидали появление дополнительных компенсаторных мутаций у вирусов из бляшек. Определение нуклеотидной последовательности, соответствующей участкам *oriL* и 3C, у этих вирусов выявило изменение последовательности тетрапетли на GAUA. Замен в 3C обнаружено не было. Мы реконструировали тетрапетлю GAUA в окружении последовательности, соответствующей таковой для генома полиовируса штамма Mahoney. Полученный вариант имел среднебляшечный фенотип на третий день после трансфекции и удельную инфекционность, близкую к таковой для контрольного транскрипта (Рисунок 27). Таким образом, замена всего лишь одного нуклеотида G на A в третьем положении тетрапетли восстанавливает функциональность репликативного элемента *oriL*.



Рисунок 27. Бляшечный фенотип вариантов, содержащих тетрапетлю GCUA с разными фланкирующими парами, тетрапетлю auGAUAgu и тетрапетлю auGGCAau, на третий день после трансфекции.

## Характеристика варианта с тетрапетлей пространственного структурного класса gCUUGc и его ревертантов

В качестве представителя еще одного не представленного в нашей выборке пространственного структурного класса тетрапетель мы сконструировали вариант, содержащий тетрапетлю gCUUGc, представителя структурного пространственного класса gCUUGc (Jucker & Pardi, 1995). Сконструированный вариант имел гетерогенный фенотип с преобладанием мелких бляшек на третий день после трансфекции культуры клеток *Vero* и удельную инфекционность одного порядка с контрольным транскриптом (Рисунок 24, Таблица 11). Таким образом, тетрапетля

в апикальном участке домена d, принадлежащая структурному классу gCUUGc, оказывает негативное влияние на жизнеспособность.

Гетерогенность бляшечного фенотипа может свидетельствовать о появлении компенсаторных мутаций. Для того чтобы выяснить, как будет меняться вариант с тетрапетлей gCUUGc, мы определили нуклеотидную последовательность участков генома, соответствующих *oriL* и белку 3CD, у вирусов из 5-ти различных бляшек. Как видно из таблицы 11, замен в области, кодирующей белок 3C, обнаружено не было. При этом последовательность тетрапетли менялась на gUUUGc или gCUCGc (Рисунок 24, Таблица 11). Варианты с тетрапетлей gUUUGc в апикальном участке домена d были обнаружены при поиске жизнеспособных из набора рандомизированных по данному участку. Соответствующий реконструированный вариант полиовируса демонстрировал фенотип, неотличимый от такового для контрольного транскрипта (Рисунок 24, Таблица 11).

Фланкирующая пара играет значительную роль в формировании специфической пространственной структуры тетрапетли gCUUGc: водородными связями с фланкирующей парой стабилизируется основание второго нуклеотида тетрапетли в малом желобке (подробнее см. обзор литературы) (Jucker & Pardi, 1995). Поэтому мы предположили, что если заменить фланкирующие пары при тетрапетле CUUG на соответствующие нуклеотидной последовательности генома вируса полиомиелита дикого типа (5'au-gu3'), то пространственная структура класса gCUUGc не должна сохраниться, хотя последовательность тетрапетли останется неизменной. Сконструированный вариант с октануклеотидом auCUUGgu после трансфекции имел фенотип и удельную инфекционность, близкую к таковой контрольного транскрипта (Рисунок 24, Таблица 11). Мы дополнительно сконструировали вариант с фланкирующей парой д...с в окружении последовательности генома полиовируса штамма Mahoney (gUGCGc) и показали, что такой вариант имеет бляшечный фенотип, неотличимый от такового для контрольного транскрипта (Рисунок 24, Таблица 11). Вместе с тем, что бляшечный фенотип и удельная инфекционность варианта с последовательностью gUUUGc в апикальном участке домена d мало отличается от таковых для контрольного транскрипта, можно сделать вывод, что фланкирующая пара д...с сама по себе не оказывает негативного влияния на жизнеспособность вируса. Мы делаем заключение, что именно структурообразующая роль фланкирующей пары, а не ее последовательность, имеет значение в случае тетрапетли CUUG. Так как при замене фланкирующих пар последовательность тетрапетли CUUG оставалась неизменной, то мы можем сделать вывод, что наблюдаемый эффект обусловлен именно ее пространственной структурой.

В совокупности, полученные данные позволяют нам сделать вывод, что способность

тетрапетли в апикальном участке домена d поддерживать эффективную репродукцию вируса определяется ее пространственной структурой, а не последовательностью.

# Характеристика вариантов полиовируса с тетрапетлями GAGA и gCUUGc и компенсаторными аминокислотными заменами в 3C.

Нарушения структуры апикального участка домена d могут быть компенсированы аминокислотными заменами Thr<sub>154</sub>Ile и Lys<sub>156</sub>Arg в 3C (Andino *et al.*, 1990b). В наших экспериментах мы также обнаруживали аминокислотную замену Thr<sub>154</sub>Ile у вариантов с неблагоприятной структурой домена d (Рисунок 28,

Таблица 12). Эта замена появлялась у вирусов с дестабилизированной (uuGCUAaa) и/или не-YNMG последовательностью (auGCACau, acCUGAgu) в апикальном участке домена d (Рисунок 28,

Таблица 12). Однако мы ни разу не обнаруживали эти компенсаторные мутации у вариантов, содержащих тетрапетли не-UNCG структурных классов - GNRA и gCUUGc. Возможно, дестабилизированные структуры, такие как uuGCUAaa, могут быть эффективно скомпенсированы заменами в 3C, а формирующие стабильную пространственную структуру классов GNRA и gCUUGc, не могут быть эффективно компенсированы этой аминокислотной заменой. Мы сконструировали геномы, содержащие тетрапетли классов GNRA или gCUUGc и аминокислотные замены Thr<sub>154</sub>Ile или Lys<sub>156</sub>Arg в 3C и сравнили их бляшечный фенотип и удельную инфекционность с таковыми для контрольного транскрипта. В качестве примера эффективной компенсации мы создали вариант, содержащий октануклеотид uuGCUAaa и компенсаторную замену Thr<sub>154</sub>Ile или Lys<sub>156</sub>Arg. Бляшечные фенотипы вариантов на четвертый день после трансфекции представлены на рисунке 29.

Вариант с тетрапетлей uuGCUAaa и аминокислотной заменой Thr<sub>154</sub>Ile имеет крупные бляшки на четвертый день после трансфекции. Таким образом, мы наблюдаем эффективную компенсацию дестабилизированной структуры uuGCUAaa аминокислотной заменой в 3С. Для вариантов с тетрапетлями GAGA и gCUUGc наличие аминокислотной замены треонина в положении 154 на изолейцин также приводит к увеличению размера бляшек (Рисунок 28,

Таблица 12). Аминокислотная замена лизина в положении 156 белка 3С на аргинин, в отличие от замены треонина в положении 154 на изолейцин, приводит к выраженному компенсаторному эффекту для варианта с тетрапетлей gCUUGc и менее выраженному для вариантов с тетрапетлей uuGCUAaa и GAGA (Рисунок 28).



Рисунок 28. Бляшечные фенотипы вариантов полиовируса с заменами в апикальном участке домена d *oriL* и аминокислотными заменами в белке 3С на 4 день после трансфекции.

Таблица 12 Бляшечный фенотип, удельная инфекционность и генетическая стабильность вариантов с мутациями в тетрапетле и белке 3С

Ν	Последовательность тетрапетли и АК мотива 154-156 в 3С	Бляшечный фенотип на 4 день	Относительная удельная инфекционность	Структура
1	GAGA	Поздние мелкие бляшки	0,004	GNRA
2	GAGA_IGK	Крупнобляшечный	0,84	GNRA
3	GAGA_TGR	Мелкобляшечный	0,17	GNRA
4	uuGCUAaa	Мелкобляшечный	0,94	Неизвестно
5	uuGCUAaa_IGK	Крупнобляшечный	0,2	Неизвестно
6	uuGCUAaa_TGR	Мелкобляшечный	0,74	Неизвестно
7	gCUUGc	Гетерогенный	0,22	gCUUGc
8	gCUUGc_IGK	Крупноблягечный	0,67	gCUUGc
9	gCUUGc_TGR	Крупнобляшечный	1,04	gCUUGc

Таким образом, мы показали, что аминокислотная замена Thr<sub>154</sub>Ile восстанавливает эффективность репродукции вариантов, содержащих тетрапетли структурных классов GNRA и gCUUGc.

В результате, мы можем сделать заключение, что тетрапетли, принадлежащие пространственному структурному классу тетрапетель UNCG, способны поддерживать эффективную репродукцию вируса полиомиелита. Тетрапетли пространственных структурных классов GNRA и gCUUGc не способны поддерживать эффективную репродукцию вируса, что и объясняет отсутствие представителей этих классов среди отобранных в эксперименте и природных вариантов полиовируса. При этом вирус полиомиелита даже при наличии не благоприятных последовательностей в апикальном участке домена d сохраняет способность к репликации и приобретает компенсаторные замены либо в самом апикальном участке домена d, либо в белке 3С.

## 3.3 Эффективность синтеза дочерних копий РНК у сконструированных вариантов

Мы предполагали, что наблюдаемые фенотипы исследованных вариантов обусловлены эффективностью взаимодействия oriL с вирусным белком 3CD, и, соответственно, эффективностью репликации. Чтобы проверить эту гипотезу, мы исследовали кинетику накопления вирусной РНК для различных вирусных вариантов (Рисунок 29). Для этого транскрипты, содержащие соответствующие мутации, были использованы для трансфекции культуры клеток *Vero*. Количество копий вирусной (+)РНК определяли в тотальной РНК, выделенной из трансфицированной культуры клеток на 0, 10, 12, 16, 20 и 24 часа после трансфекции.

Эффективность репликации исследуемых вариантов коррелировала с их бляшечным фенотипом, что позволяет нам сделать заключение, что эффективность репродукции мутантов обусловлена именно этим этапом вирусного цикла.

Так, варианты, имеющие крупнобляшечный фенотип (uGCUAg, uUGAGg, uUGUGg, uGGUAg, uCUUGg), демонстрируют высокую эффективность накопления дочерних копий PHK, сравнимую с таковой для контрольного транскрипта (Рисунок 29, А, Б, В).

Представители пространственных структурных классов тетрапетель GNRA (GAGA) и gCUUGc, напротив, не могли поддерживать эффективную репликацию (Рисунок 29 Б, В). Вариант, содержащий тетрапетлю GAUA – тетрапетля псевдоревертанта полиовируса с тетрапетлей GAGA, – реплицировался значительно эффективнее, чем полиовирус с тетрапетлей GAGA (Рисунок 29 В). Геном, содержащий тетрапетлю CUUG с прилегающей спиралью дикого типа au-gu, в отличие от

генома с октануклеотидом agCUUGcu, реплицировался на уровне контрольного транскрипта (Рисунок 29 Б).



Рисунок 29. Кинетика накопления дочерних вирусных геномов. В каждой точке указано среднеквадратичное отклонение. А –эффективность репликации вариантов с uUGUGg, uUGAGg и uGGUAg последовательностями; Б – эффективность репликации вариантов с gCUUGc и uCUUGg последовательностями; В - эффективность репликации вариантов с пространственными структурами GNRA (GAGA) и UNCG (GCUA, GAUA) в апикальном участке домена d.

Анализ кинетики накопления дочерних РНК для варианта, содержащего тетрапетлю GAGA и аминокислотный триплет Thr<sub>154</sub>Ile в белке 3C, показал, что исследуемая аминокислотная замена частично компенсирует неспособность генома с тетрапетлей GAGA к репликации (Рисунок 30).



Рисунок 30. Сравнение кинетики синтеза РНК геномов GAGA с триплетами TGK и IGK с таковой для контрольного транскрипта.

## 3.4 Влияние пространственной структуры апикального участка домена d oriL на эффективность его взаимодействия с белком 3CD

Для оценки эффективности взаимодействия различных мутантных oriL с рекомбинантным белком 3CD использовали РНК-транскрипты, соответствующий первым 113 положениям генома полиовируса с внесенными мутациями. Визуализацию комплексов после ЭФ в не денатурирующем 6%-ном ПААГ осуществляли двумя методами: 1) окраска РНК бромистым этидием и 2) окраска белка 3CD методом Western-blot с использованием антител к препарату рекомбинантного 3CD, Мы наблюдали образование двух комплексов с различной полученных в мышах. электрофоретической подвижностью (Рисунок 31). Чтобы выяснить, не является ли какой-либо из наблюдаемых комплексов неспецифическим комплексом с белками E.coli, был поставлен аналогичный эксперимент с лизатом E.coli, трансформированным пустым вектором pQE60. Был приготовлен лизат и проведены аналогичные стадии очистки, как и при выделении рекомбинантного 3CD. В экспериментах с таким лизатом наблюдалось эффективное формирование нижнего комплекса без формирования верхнего комплекса. Таким образом, нижний комплекс является неспецифическим комплексом с соочищающимся белком E.coli. Чтобы подтвердить, что верхний 90

комплекс является специфическим, были поставлены дополнительные эксперименты. Верхний и нижний комплексы были вырезаны и полученные куски геля были помещены в лунки денатурирующего ПААГ для электрофореза по Лэммли. По окончании электрофореза методом Вестерн-блот с антителами к препарату рекомбинантного 3CD в соответствующих дорожках был обнаружен белок, подвижность которого соответствовала 3CD (Рисунок 31). В аналогичных условиях в экспериментах с oriL, содержащим в апикальном участке домена d тетрапетлю GAGA, 3CD обнаружено не было (Рисунок 32).

Наличие 3CD в вырезанном из геля комплексе было дополнительно подтверждено методом MALDI—MS (масс-спектрометрия выполнена Серебряковой М.В. в НИИ физико-химической биологии им. А.Н. Белозерского МГУ).



100% 0% 23% 42% 4% 73% 2% 0%

Рисунок 31. Электрофореграмма эксперимента EMSA (Electrophpretical mobility shift assay). А - окраска бромистым этидием. Б – EMSA и последующим Western-blot с сывороткой, полученной в мышах к препарату рекомбинантного 3CD. Цифры под панелью Б – относительное количество комплекса, определенное денситометрически.



Рисунок 32. Электрофорез и последующий Western Blot белков, содержащихся в вырезанных после EMSA комплексах. Для эксперимента были использованы препараты 3CD дикого типа или с аминокислотной заменой Lys<sub>156</sub>Met (соответствующий белок 3CD не имеет PHK-связывающей активности, не опубликовано) и препараты *oriL* с тетрапетлей дикого типа либо с тетрапетлей GAGA.

Эффективность взаимодействия oriL с различными тетрапетлями в апикальном участке домена d коррелирует с эффективностью репликации соответствующих вариантов и их фенотипом. В этих экспериментах oriL, содержащий тетрапетли с пространственными структурами GNRA или gCUUGc, практически не связывается с рекомбинантным белком 3CD (Рисунок 31, дорожки 2 и 5;). Напротив, oriL, содержащая тетрапетлю, имеющую свойства структурного класса UNCG – GCUA - связывается с 3CD, хотя и менее эффективно, чем контрольный oriL (Рисунок 31, дорожка 4). Эффективность связывания резко падает при введении фланкирующей пары uu-aa при тетрапетле GCUA (Рисунок 31, дорожка 7). Замена фланкирующей пары при тетрапетле CUUG с g-с на u-g восстанавливает эффективность связывания с 3CD (Рисунок 31, дорожка 6). Также, oriL, содержащая выпетленный пятый нуклеотид в апикальном участке домена d, не узнается рекомбинантным 3CD (Рисунок 31, дорожка 8). Эксперименты были повторены два раза (приведены данные одного опыта) с независимо полученными препаратами PHK и 3CD с одинаковым результатом. Таким образом, наблюдаемые бляшечные фенотипы и эффективность репликации обусловлены эффективностью взаимодействия 3CD и oriL.

### 4. Обсуждение

Из набора вариантов, содержащих случайную последовательность вместо 8 апикальных нуклеотидов домена d oriL, мы получили 39 уникальных жизнеспособных нуклеотидных последовательности. Анализ полученных последовательностей показал, что практически в любом положении исследуемого октануклеотида может быть любой нуклеотидный остаток. Единственное исключение составляло шестое положение октонуклеотида (четвертое положение тетрапетли), в котором ни разу не был обнаружен U. Несмотря на то, что мы не обнаружили консервации определенных нуклеотидных остатков в каких-либо положениях исследуемого октонуклеотида, удалось обнаружить корреляцию между нуклеотидными остатками в разных положениях и, на их основе, разбить 33 уникальных последовательности из 39 на три консенсуса. Этими консенсусами стали YNMG (Y=C/U, N=любой, M=A/C), YNUG и GSYA (S=G/C). Реконструкция вариантов из каждой группы в контексте последовательности генома полиовируса дикого типа (штамм Mahoney) и их вирусологическая характеристика показала, что все они способны поддерживать репродукцию полиовируса, несмотря на существенную разницу в последовательности апикального участка. Следовательно, мы можем сделать заключение, что нуклеотидная последовательность апикального участка домена d не играет решающей роли. Реконструкция дополнительных вариантов, не отобранных экспериментально, позволила уточнить пространство последовательностей тетрапетель в апикальном участке домена d, способных поддерживать эффективную репродукцию полиовируса. Еще 6 последовательностей, которые нельзя было отнести к каким-либо консенсусам. Оказалось, что эти последовательности не могут эффективно поддерживать репродукцию полиовируса и в процессе пассирования приобретают мутации в апикальном участке домена d, приводящие ее к одному из трех описанных выше консенсусов, или компенсируются аминокислотными заменами в 3С. Таким образом, множество вариантов тетрапетель, способных поддерживать эффективную репродукцию вируса, можно описать тремя консенсусами nYNMGm, nYNUGm (кроме gCUUGc) и uGVUAg, где n и m – любая комплиментарная пара нуклеотидов или и и g, M=A/C, V - все нуклеотиды, кроме уридина.

На основании наших данных мы можем оценить, какова будет вероятность сохранения фитнеса, сравнимого с диким типом, в случае одной мутации в тетрапетле, то есть уровень помехоустойчивости этого элемента. Для примера возьмем тетрапетлю полиовируса первого типа штамма Mahoney UGCG. Условия сохранения фенотипа дикого типа по нашим данным: первое положение занимает пиримидин, второе - любой нуклеотид, третье - не гуанозин, четвертое только

гуанозин. Всего возможно 12 вариантов: по три замены на каждый нуклеотид. С учетом вышеописанных условий, 6 из них (50%) будут репродуцироваться с эффективностью полиовируса штамма Mahoney (CGCG, UACG, UCCG, UUCG, UGAG, UGUG). При этом любой из 12 вариантов будет жизнеспособным.

Рассмотрим, какова будет вероятность сохранения высокой эффективности репродукции в случае одной мутации в октануклеотиде. Всего возможно 24 варианта: по три возможных замены в восьми положениях. Из них 12 – замены в тетрапетле, из которых 6 продуцируют вирус с крупнобляшечным фенотипом, как показано выше. Другие 12 - замены во фланкирующих парах. Мы показали, что в случае одной не спаренной пары в спирали (не а-u, g-c или u-g), вирус будет достаточно эффективно репродуцироваться. Таким образом, условия сохранения жизнеспособности позволяют одну разрушенную фланкирующую пару, то есть любую замену в этом участке. Итого: 6+12 вариантов=18 вариантов. То есть, вероятность сохранения высокого фитнеса в случае одной замены в октануклеотиде очень высока - 75%.

Вирусная полимераза при синтезе дочерних копий вносит около одной замены на каждый дочерний геном, что позволяет вирусу быстро приспособиться к изменяющимся условиям среды (Acevedo, Brodsky & Andino, 2014). При этом трансмиссия вируса сопряжена с эффектом бутылочного горлышка, когда сохранение популяции может зависеть от жизнеспособности нескольких вирусных частиц. В свете этих двух противоположенных процессов, механизмы, позволяющие вирусу сохранять функциональность генома при появлении в нем мутаций, то есть механизмы помехоустойчивости, могут подвергаться положительному отбору. Например, в ходе цикла репродукции вируса упаковываются, в основном, дочерние геномы пятого поколения, то есть те, которые уже продемонстрировали способность к репликации (Schulte *et al.*, 2015). Таким образом, мутации в апикальном участке домена d, снижающие способность генома к репликации будут подвергаться очищающему отбору еще до упаковки дочерних геномов в капсид. Это позволяет сохранить одновременно высокую скорость появления мутаций и функциональность генома, то есть, является одним из механизмов помехоустойчивости (Schulte *et al.*, 2015). Наши данные демонстрируют высокий уровень помехоустойчивости апикального участка домена d: ни одна из возможных замен в нем не является летальной для вируса.

С другой стороны, известно, что у близких родственников полиовируса апикальная тетрапетля и ее прилегающая фланкирующая пара напрямую вовлечены во взаимодействие с белком 3CD и что это взаимодействие является одним из ключевых событий инициации репликации вирусного генома (Andino *et al.*, 1993; Gamarnik & Andino, 1998; Ohlenschläger *et al.*, 2004; Ihle *et al.*,

2005). Эти два факта предполагают, что последовательность этого элемента должна быть консервативна, что противоречит нашим результатам.

Мы считаем, что одной из причин, почему этот элемент сохраняет свою функциональность при значительных изменениях своей последовательности является то, что белок 3CD узнает скорее определенную пространственную структуру этого участка, а не специфическую последовательность нуклеотидных остатков.

Ранее было показано, что исследованные тетрапетли домена d разных энтеровирусов (CVB3, HRV2 и BEV1) имеют пространственную структуру класса UNCG (Du *et al.*, 2003, 2004; Ihle *et al.*, 2005). Для тех тетрапетель, которые были обнаружены у вирусов, отобранных в наших экспериментах из набора со случайными последовательностями и способных к эффективной репродукции, для которых известна пространственная структура, она также относится к классу UNCG (UGCG, UACG, UCAG, CCCG, UUCG, GCUA, UUUG) (Varani *et al.*, 1991; Abdelkafi *et al.*, 1998; Finger *et al.*, 2003; Du *et al.*, 2004; Melchers *et al.*, 2006; Zhou *et al.*, 2007; Miller *et al.*, 2014).

Для того, чтобы оценить, могут ли тетрапетли с пространственной структурой, отличной от структуры класса UNCG, поддерживать эффективную репликацию вирусного генома, мы сконструировали геномы, содержащие тетрапетли классов GNRA и gCUUGc в апикальном участке домена d. Тетрапетли в апикальном участке домена d, принадлежащие к пространственным классам GNRA и gCUUGc, не способны поддерживать эффективную репликацию вирусного генома, хотя и не являются летальными и могут быть компенсированы заменами в белке 3С. В пользу того, что негативный эффект на репродукцию полиовируса оказывает именно пространственная структура, а не последовательность этих тетрапетель, говорит тот факт, что замена фланкирующих пар при тетрапетле CUUG с g-с на u-g возвращает тетрапетле CUUG способность поддерживать вирусную репродукцию. При том, что сама по себе фланкирующая пара g-с не оказывает негативного эффекта, это свидетельствует о том, что именно пространственная структура, стабилизированная фланкирующей парой g-с, а не последовательность тетрапетли CUUG оказывает негативный эффект на вирусную репликацию.

Элемент oriL участвует в ряде других процессов, кроме инициации синтеза (-) цепи, поэтому вносимые нами мутации, могли сказываться на других этапах репликации вирусного генома. Корреляция эффективности репликации мутантных вирусных геномов и эффективности взаимодействия соответствующих oriL с рекомбинантным 3CD свидетельствует о том, что наблюдаемые эффекты связаны именно со взаимодействием oriL с вирусным белком 3CD.

Таким образом, наши данные подтверждают гипотезу о том, что пространственная структура

апикальной тетрапетли домена d более значима для узнавания вирусным белком 3CD, чем ее последовательность. Пространственная структура класса UNCG может быть реализована очень широким набором нуклеотидных последовательностей (Petrov, Zirbel & Leontis, 2013). И, таким образом, наличие именно этой пространственной структуры определяет высокий уровень помехоустойчивости этого элемента.

Можно предположить, что тетрапетли, обнаруженные у вирусов, не способных к эффективной репродукции, формируют пространственную структуру, отличную от таковой для UNCG. Например, вирус, содержащий тетрапетлю AGCA в апикальном участке домена d, не способен к эффективной репродукции и с пассажами приобретает замены в тетрапетле и аминокислотную замену Thr<sub>154</sub>IIe в белке 3C, известную, как компенсаторная. При этом существуют косвенные данные, указывающие на то, что тетрапетля AGCA предположительно имеет структуру пространственного класса AGNN, которая отличается от таковой класса UNCG (Chanfreau, Buckle & Jacquier, 2000; Lebars *et al.*, 2001; Wu *et al.*, 2001, 2004; Wang *et al.*, 2011).

Чтобы понять более детально, как осуществляется узнавание апикального участка домена d белком 3CD, мы сконструировали вирус, содержащий в этом участке пентапетлю UCAGU, первые четыре нуклеотида которой формируют пространственную структуру класса UNCG, а основание пятого нуклеотида выпетлено в большой желобок с 3'-стороны тетрапетли (Theimer *et al.*, 2003). Такой вариант не был способен к эффективной репродукции и с пассажами приобретал аминокислотные замены в 3C, в том числе замену Thr<sub>154</sub>Ile, известную, как компенсаторная. Соответствующий oriL не узнавался белком 3CD. Таким образом, выпетленный пятый нуклеотид создает стерические затруднения для узнавания белком, что указывает на то, что узнавание апикального участка домена d, хотя бы отчасти, осуществляется со стороны большого желобка тетрапетли.

Рассмотрим, какие структурные элементы со стороны большого желобка апикального участка домена d могут быть вовлечены во взаимодействие с 3CD.

Структурным элементом, который может узнаваться белком 3CD с 3'-стороны большого желобка тетрапетли и характерен для всех тетрапетель пространственного класса UNCG, является так называемый Z-step (от Z-ДНК) (D'Ascenzo *et al.*, 2016). В данном случае этот структурный элемент характеризуется специфическим π-взаимодействием и ориентацией основания четвертого нуклеотида и рибозой третьего нуклеотида (D'Ascenzo *et al.*, 2016).

Еще одним элементом, вовлеченным в узнавание, могут быть фосфаты сахарофосфатного остова, который имеет специфическую структуру у тетрапетель класса UNCG. Известно, что в PHK-

белковом узнавании половина взаимодействующих группировок приходится на фосфаты сахарофосфатного остова (Jones *et al.*, 2001; Bahadur, Zacharias & Janin, 2008). Со стороны 3С кандидатами для взаимодействия с фосфатами сахарофосфатного остова могут быть N-концевая спираль, несущая вовлеченные в узнавание аминокислоты  $Lys_{12}$  и  $Arg_{13}$  и/или линкер между бочонками, содержащий консервативную последовательность  $s_2$ KFRDI<sub>86</sub> (Blair *et al.*, 1998; Amero *et al.*, 2008). И тот и другой участки PHK-связывающего сайта богаты положительно заряженными аргинином и лизином, для которых характерно взаимодействие с фосфатами сахарофосфатного остова в PHK-белковых комплексах (Jones *et al.*, 2001).

Единственная совпадающая химическая группа у благоприятных для репродукции полиовируса тетрапетель UACG, CACG, GCUA и отличающая их от не благоприятных для репродукции тетрапетель gCUUGc и GNRA (анализ по данным ЯМР, см. Таблица 13) – это аминогруппа (-NH<sub>2</sub>) с 3' стороны тетрапетли (Таблица 13). У изученных тетрапетель, принадлежащих консенсусу YNMG, эта группа представлена остатком четвертого гуанозина, а у тетрапетли GCUA остатком первого гуанозина. Тетрапетля GNRA практически не имеет в большом желобке реакционных групп для взаимодействия. Обсуждаемая аминогруппа у изученных тетрапетель, принадлежащих консенсусу YNMG, расположена довольно глубоко в большом желобке, и, без дополнительного раскрытия структуры, взаимодействие с ней может быть затруднено. Дополнительное раскрытие большого желобка может осуществляться при расплетании домена d, так как это происходит при взаимодействии 3CD и репликативного элемента *ori*I (Shen *et al.*, 2007; Pathak *et al.*, 2007).

	UACG	CACG	GCUA	gCUUGc
	Крупно-	Крупно-	Крупно-	Мелко-
Фенотип	бляшечный	бляшечный	бляшечный	бляшечный
Группы в положениях 1 и 4 тетрапетли	(4)NH <sub>2</sub> -(1)O	(4)NH <sub>2</sub> -(1)NH <sub>2</sub>	(4)NH <sub>2</sub> -(1)O	(4)O-(1) NH <sub>2</sub>
Группы в положении 3 тетрапетли	(3)O-(3)NH <sub>2</sub>	(3)O-(3)NH <sub>2</sub>	(3)0-(3)0	(3)0-(3)0

Таблица 13 Сравнение реакционных групп, представленных в большом желобке, у различных тетрапетель (оценка по визуализации файлов из базы данных PDB в программе VMD)

Данные литературы и наши результаты показывают, что неблагоприятные замены в

апикальном участке домена d, как в тетрапетле, так и в прилегающей спирали, могут компенсироваться аминокислотными заменами в РНК-связывающем сайте в 3С (Andino et al., 1990b). В наших экспериментах, это, в основном, замена полярного треонина в положении 154 на гидрофобный изолейцин, реже замена лизина на аргинин с более высоким положительным зарядом в положении 156. Замена треонина в положении 154 ЗС на изолейцин представляется универсальной практически для всех мутаций в районе тетрапетли – она компенсирует такие разные нарушения, как дестабилизирующие фланкирующие пары при тетрапетле GCUA, неблагоприятные мутации в самой тетрапетле (как например, у тетрапетель AGCA и GGCA) и, возможно, выпетленный в большой желобок пятый нуклеотид при пентапетле UCAGU. Компенсация таких разных нарушений, как дестабилизация пространственной структуры тетрапетли или стерическое затруднение для узнавания при сохранении пространственной структуры тетрапетли, должно требовать разных типов изменений, если было бы задействовано прямое узнавание. Так как все эти нарушения компенсируются одной и той же аминокислотной заменой, мы предполагаем, что треонин в положении 154 не принимает участия в прямом узнавании тетрапетли, но, возможно, определяет степень конформационной подвижности элементов РНК-узнающего сайта или структуру поверхности взаимодействия.

### 5. Заключение

В данной работе продемонстрировано, что любое положение апикального октонуклеотида домена d oriL может быть занято практически любым нуклеотидным остатком (кроме шестого положения, где не может быть U). При этом апикальный участок домена d oriL способен поддерживать эффективную репликацию в широком диапазоне последовательностей, которые можно описать консенсусами: nYNHGm (за исключением gCUUGc) и uGVUAg, где n и m – любые комплементарные нуклеотиды, H – любой нуклеотид, кроме G, V – любой нуклеотид, кроме U. Такая широкая вариабельность этого функционального участка объясняется тем, что связывающий его вирусный белок 3CD, по-видимому, узнает пространственную структуру класса UNCG, которая может быть реализована большим количеством разнообразных последовательностей.

### 6. Выводы

- 1. Эффективная репродукция вируса полиомиелита с природной последовательностью 3С поддерживается множеством тетрапетель в апикальном участке домена d *oriL*, которое можно описать консенсусами: nYNHGm (за исключением gCUUGc) и uGVUAg, где n и m любые комплементарные нуклеотиды, H любой нуклеотид, кроме G, V любой нуклеотид, кроме U.
- 2. Исследованные тетрапетли в апикальном участке домена d, имеющие пространственную структуру класса UNCG, обеспечивают взаимодействие репликативного элемента oriL с вирусным белком 3CD и поддерживают эффективную репликацию вируса полиомиелита. Тетрапетли в апикальном участке домена d, имеющие пространственную структуру классов GNRA и gCUUGc, не обеспечивают взаимодействие репликативного элемента oriL с вирусным белком 3CD, и, соответственно, не могут поддерживать эффективную репликацию вируса полиомиелита. На основании этих данных может быть выдвинута гипотеза об первичном значении пространственной структуры апикального участка домена d oriL для взаимодействия с вирусным белком 3CD.
- 3. Аминокислотная замена Thr<sub>154</sub>Ile в белке 3C может компенсировать различные нарушения структуры апикального участка домена d и наличие тетрапетель, чья пространственная структура отличается от структуры класса UNCG.
- 4. Создана новая эффективная бактериальная система экспрессии и разработана методика получения (выделения и очистки) высокоэффективного вирусного рекомбинантного белка 3CD.

### 7. Список литературы

- Abdelkafi, M., Leulliot, N., Baumruk, V., Bednárová, L., Turpin, P.Y., Namane, A., Gouyette, C., Huynh-Dinh, T. & Ghomi, M. (1998). Structural features of the UCCG and UGCG tetraloops in very short hairpins as evidenced by optical spectroscopy. *Biochemistry* 37, 7878–7884.
- Acevedo, A., Brodsky, L. & Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **505**, 686–690.
- Agol, V.I., Paul, A. V. & Wimmer, E. (1999). Paradoxes of the replication of picornaviral genomes. *Virus Res.* **62**, 129–147.
- Al-Hashimi, H.M. (2013). NMR studies of nucleic acid dynamics. J. Magn. Reson. 237, 191-204.
- Aldabe, R., Barco, A. & Carrasco, L. (1996). Membrane permeabilization by poliovirus proteins 2B and 2BC. J. Biol. Chem. 271, 23134–23137.
- Aldabe, R. & Carrasco, L. (1995). Induction of membrane proliferation by poliovirus proteins 2C and 2BC. Biochem. Biophys. Res. Commun. 206, 64–76.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–402.
- Amero, C.D., Arnold, J.J., Moustafa, I.M., Cameron, C.E. & Foster, M.P. (2008). Identification of the orilbinding site of poliovirus 3C protein by nuclear magnetic resonance spectroscopy. J. Virol. 82, 4363– 4370.
- Andino, R., Rieckhof, G.E., Achacoso, P.L. & Baltimore, D. (1993). Poliovirus RNA synthesis utilizes an RNP complex formed around the 5' -end of viral RNA. *EMBO J.* **12**, 3587–3598.
- Andino, R., Rieckhof, G.E. & Baltimore, D. (1990a). A Functional Ribonucleoprotein around the 5' End of Poliovirus. *Cell* 63, 369–380.
- Andino, R., Rieckhof, G.E., Trono, D. & Baltimore, D. (1990b). Substitutions in the protease (3Cpro) gene of poliovirus can suppress a mutation in the 5' noncoding region. *J Virol* **64**, 607–612.
- Antao, V.P., Lai, S.Y. & Tinoco, I. (1991). A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.* 19, 5901–5905.
- Antao, V.P. & Tinoco, I. (1992). Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res.* **20**, 819–824.
- Asare, E., Mugavero, J., Jiang, P., Wimmer, E. & Paul, A. V. (2016). A Single Amino Acid Substitution in Poliovirus Nonstructural Protein 2C <sup>ATPase</sup> Causes Conditional Defects in Encapsidation and Uncoating. J. Virol. 90, 6174–6186.
- Bahadur, R.P., Zacharias, M. & Janin, J. (2008). Dissecting protein-RNA recognition sites. *Nucleic Acids Res.* **36**, 2705–2716.
- Baltimore, D. (1964). In vitro synthesis of viral RNA by the poliovirus RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **51**, 450–456.
- Baltimore, D., Eggers, H.J., Franklin, R.M. & Tamm, I. (1963). Poliovirus-induced RNA polymerase and the effects of virus-specific inhibitors on its production. *Proc. Natl. Acad. Sci. U. S. A.* **49**, 843–849.
- Barco, A. & Carrasco, L. (1995). A human virus protein, poliovirus protein 2BC, induces membrane proliferation and blocks the exocytic pathway in the yeast Saccharomyces cerevisiae. *EMBO J.* 14, 3349–3364.
- Barton, D.J. & Flanegan, J.B. (1997). Synchronous replication of poliovirus RNA: initiation of negativestrand RNA synthesis requires the guanidine-inhibited activity of protein 2C. J. Virol. 71, 8482–8489.
- Barton, D.J., Morasco, B.J. & Flanegan, J.B. (1999). Translating ribosomes inhibit poliovirus negativestrand RNA synthesis. J. Virol. 73, 10104–10112.
- Barton, D.J., O'Donnell, B.J. & Flanegan, J.B. (2001). 5' cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. *EMBO J.* **20**, 1439–1448.

- Basavappa, R., Filman, D.J., Syed, R., Flore, O., Icenogle, J.P. & Hogle, J.M. (1994). Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: Structure of the empty capsid assembly intermediate at 2.9 Å resolution. *Protein Sci.* **3**, 1651–1669.
- Bedard, K.M., Daijogo, S. & Semler, B.L. (2007). A nucleo-cytoplasmic SR protein functions in viral IRESmediated translation initiation. *EMBO J.* 26, 459–467.
- Bellmunt, A., May, G., Zell, R., Pring-Åkerblom, P., Verhagen, W. & Heim, A. (1999). Evolution of Poliovirus Type I during 5.5 Years of Prolonged Enteral Replication in an Immunodeficient Patient. *Virology* 265, 178–184.
- Belov, G.A., Altan-Bonnet, N., Kovtunovych, G., Jackson, C.L., Lippincott-Schwartz, J. & Ehrenfeld, E. (2007). Hijacking components of the cellular secretory pathway for replication of poliovirus RNA. J. Virol. 81, 558–567.
- Belov, G.A. & Ehrenfeld, E. (2007). Involvement of cellular membrane traffic proteins in poliovirus replication. *Cell cycle* **6**, 36–38.
- Belov, G.A., Nair, V., Hansen, B.T., Hoyt, F.H., Fischer, E.R. & Ehrenfeld, E. (2012). Complex dynamic development of poliovirus membranous replication complexes. J. Virol. 86, 302–312.
- Bernstein, H.D., Sarnow, P. & Baltimore, D. (1986). Genetic complementation among poliovirus mutants derived from an infectious cDNA clone. J. Virol. 60, 1040–1049.
- Bessaud, M., Joffret, M.-L., Blondel, B. & Delpeyroux, F. (2016). Exchanges of genomic domains between poliovirus and other cocirculating species C enteroviruses reveal a high degree of plasticity. *Sci. Rep.* 6, 38831.
- Bienz, K., Egger, D. & Pasamontes, L. (1987). Association of polioviral proteins of the P2 genomic region with the viral replication complex and virus-induced membrane synthesis as visualized by electron microscopic immunocytochemistry and autoradiography. *Virology* **160**, 220–226.
- Bienz, K., Egger, D., Troxler, M. & Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. J. Virol. 64, 1156–1163.
- Blair, W.S., Nguyen, J.H., Parsley, T.B. & Semler, B.L. (1996). Mutations in the poliovirus 3CD proteinase S1-specificity pocket affect substrate recognition and RNA binding. *Virology* **218**, 1–13.
- Blair, W.S., Parsley, T.B., Bogerd, H.P., Towner, J.S., Semler, B.L. & Cullen, B.R. (1998). Utilization of a mammalian cell-based RNA binding assay to characterize the RNA binding properties of picornavirus 3C proteinases. *RNA* **4**, 215–25.
- Blyn, L.B., Swiderek, K.M., Richards, O., Stahl, D.C., Semler, B.L. & Ehrenfeld, E. (1996). Poly( rC ) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5 B<sup>TM</sup> noncoding region: Identification by automated liquid chromatography-tandem mass spectrometry. *Proc Natl Acad Sci USA*. **93**, 11115–11120.
- Blyn, L.B., Towner, J.S., Semler, B.L. & Ehrenfeld, E. (1997). Requirement of poly(rC) binding protein 2 for translation of poliovirus RNA. *J. Virol.* **71**, 6243–6246.
- Bolten, R., Egger, D., Gosert, R., Schaub, G., Landmann, L. & Bienz, K. (1998). Intracellular localization of poliovirus plus- and minus-strand RNA visualized by strand-specific fluorescent In situ hybridization. J. Virol. 72, 8578–85.
- Bonderoff, J.M., Larey, J.L. & Lloyd, R.E. (2008). Cleavage of Poly (A) -Binding Protein by Poliovirus 3C Proteinase Inhibits Viral Internal Ribosome Entry Site-Mediated Translation. J. Virol. 82, 9389–9399.
- Bothe, J.R., Nikolova, E.N., Eichhorn, C.D., Chugh, J., Hansen, A.L. & Al-Hashimi, H.M. (2011). Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy. *Nat. Methods* **8**, 919–931.
- Bottaro, S., Banáš, P., Šponer, J. & Bussi, G. (2016a). Free Energy Landscape of GAGA and UUCG RNA Tetraloops. J. Phys. Chem. Lett. 7, 4032–4038.
- Bottaro, S., Gil-Ley, A. & Bussi, G. (2016b). RNA folding pathways in stop motion. Nucleic Acids Res. 44,

5883-5891.

- Bottaro, S. & Lindorff-Larsen, K. (2017). Mapping the Universe of RNA Tetraloop Folds. *Biophys. J.* **113**, 257–267.
- Bottaro, S., Di Palma, F. & Bussi, G. (2014). The role of nucleobase interactions in RNA structure and dynamics. *Nucleic Acids Res.* **42**, 13306–14.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–54.
- Brown, D.M., Cornell, C.T., Tran, G.P., Nguyen, J.H.C. & Semler, B.L. (2005). An authentic 3' noncoding region is necessary for efficient poliovirus replication. *J. Virol.* **79**, 11962–11973.
- Brunner, J.E., Nguyen, J.H.C., Roehl, H.H., Ho, T. V, Swiderek, K.M. & Semler, B.L. (2005). Functional Interaction of Heterogeneous Nuclear Ribonucleoprotein C with Poliovirus RNA Synthesis Initiation Complexes. J. Virol. 79, 3254–3266.
- Burns, C.C., Lawson, M.A., Semler, B.L. & Ehrenfeld, E. (1989). Effects of mutations in poliovirus 3Dpol on RNA polymerase activity and on polyprotein cleavage. J. Virol. 63, 4866–4874.
- Burrill, C.P., Westesson, O., Schulte, M.B., Strings, V.R., Segal, M. & Andino, R. (2013). Global RNA Structure Analysis of Poliovirus Identifies Conserved RNA Structure Involved in Viral Replication and Infectivity. J. Virol. 87, 11670–11683.
- Butcher, S.E., Dieckmann, T. & Feigon, J. (1997). Solution structure of the conserved 16 S-like ribosomal RNA UGAA tetraloop. *J. Mol. Biol.* **268**, 348–358.
- Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A. & Patel, D.J. (1998). Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. *Nat. Struct. Biol.* 5, 203–212.
- Cameron, C.E., Oh, H.S. & Moustafa, I.M. (2010). Expanding knowledge of P3 proteins in the poliovirus lifecycle. *Future Microbiol.* **5**, 867–881.
- Castello, A., Alvarez, E. & Carrasco, L. (2006). Differential Cleavage of eIF4GI and eIF4GII in Mammalian Cells: effects on translation. *J. Biol. Chem.* **281**, 33206–33216.
- Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R. & Doudna, J.A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science (80-. )*. 273, 1678–1685.
- Chanfreau, G., Buckle, M. & Jacquier, A. (2000). Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. *Proc. Natl. Acad. Sci. U. S. A.* 97, 3142–3147.
- Chase, A.J., Daijogo, S. & Semler, B.L. (2014). Inhibition of poliovirus-induced cleavage of cellular protein PCBP2 reduces the levels of viral RNA replication. *J. Virol.* **88**, 3192–3201.
- Cheong, C. & Cheong, H. (2010). RNA Structure: Tetraloops. In *Encyclopedia of life sciences*. Chichester: John Wiley & Sons, Ltd.
- Cheong, C., Varani, G. & Tinoco, I.J. (1990). Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature*. **346**, 680–682.
- Cho, M.W., Teterina, N., Egger, D., Bienz, K. & Ehrenfeld, E. (1994). Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. *Virology* **202**, 129–145.
- Claridge, J.K., Headey, S.J., Chow, J.Y.H., Schwalbe, M., Edwards, P.J., Jeffries, C.M., Venugopal, H., Trewhella, J. & Pascal, S.M. (2009). A picornaviral loop-to-loop replication complex. *J. Struct. Biol.* **166**, 251–262.
- Collis, P.S., O'Donnell, B.J., Barton, D.J., Rogers, J.A. & Flanegan, J.B. (1992). Replication of poliovirus RNA and subgenomic RNA transcripts in transfected cells. *J. Virol.* **66**, 6480–6488.
- Cornell, C.T. & Semler, B.L. (2002). Subdomain Specific Functions of the RNA Polymerase Region of Poliovirus 3CD Polypeptide. *Virology* 298, 200–213.
- Correll, C.C. & Swinger, K. (2003). Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 A resolution. *RNA* 9, 355–363.

- Correll, C.C., Wool, I.G. & Munishkin, A. (1999). The two faces of the Escherichia coli 23 S rRNA sarcin/ricin domain: the structure at 1.11 A resolution. J. Mol. Biol. 292, 275–287.
- D'Ascenzo, L., Leonarski, F., Vicens, Q. & Auffinger, P. (2016). "Z-DNA like" fragments in RNA: a recurring structural motif with implications for folding, RNA/protein recognition and immune response. *Nucleic Acids Res.* 44, 5944–5956.
- Daijogo, S. & Semler, B.L. (2011). Mechanistic intersections between picornavirus translation and RNA replication. *Adv. Virus Res.* **80**, 1–24.
- Database resources of the National Center for Biotechnology Information. (2014). *Nucleic Acids Res.* **43**, D6-17.
- DeStefano, J.J. (2010). Effect of reaction conditions and 3AB on the mutation rate of poliovirus RNAdependent RNA polymerase in a alpha-complementation assay. *Virus Res.* 147, 53–59.
- Dewalt, P.G. & Semler, B.L. (1987). Site-directed mutagenesis of proteinase 3C results in a poliovirus deficient in synthesis of viral RNA polymerase. J. Virol. 61, 2162–70.
- Dorsch-Häsler, K., Yogo, Y. & Wimmer, E. (1975). Replication of picornaviruses. I. Evidence from in vitro RNA synthesis that poly(A) of the poliovirus genome is genetically coded. *J. Virol.* **16**, 1512–7.
- Drexler, J.F., Grard, G., Lukashev, A.N., Kozlovskaya, L.I., Böttcher, S., Uslu, G., Reimerink, J., Gmyl, A.P., Taty-Taty, R., Lekana-Douki, S.E., Nkoghe, D., Eis-Hübinger, A.M., Diedrich, S., Koopmans, M., Leroy, E.M. & Drosten, C. (2014). Robustness against serum neutralization of a poliovirus type 1 from a lethal epidemic of poliomyelitis in the Republic of Congo in 2010. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12889–94.
- Du, Z., Yu, J., Andino, R. & James, T.L. (2003). Extending the family of UNCG-like tetraloop motifs: NMR structure of a CACG tetraloop from coxsackievirus B3. *Biochemistry* **42**, 4373–4383.
- Du, Z., Yu, J., Ulyanov, N.B., Andino, R. & James, T.L. (2004). Solution structure of a consensus stemloop D RNA domain that plays important roles in regulating translation and replication in enteroviruses and rhinoviruses. *Biochemistry* **43**, 11959–11972.
- Egger, D. & Bienz, K. (2002). Recombination of poliovirus RNA proceeds in mixed replication complexes originating from distinct replication start sites. *J. Virol.* **76**, 10960–10971.
- Egger, D. & Bienz, K. (2005). Intracellular location and translocation of silent and active poliovirus replication complexes. J. Gen. Virol. 86, 707–718.
- Egger, D., Teterina, N., Ehrenfeld, E. & Bienz, K. (2000). Formation of the poliovirus replication complex requires coupled viral translation, vesicle production, and viral RNA synthesis. *J Virol* **74**, 6570–6580.
- Eggers, H.J. (1999). Milestones in early poliomyelitis research (1840 to 1949). J. Virol. 73, 4533–4535.
- Ennifar, E., Nikulin, A., Tishchenko, S., Serganov, A., Nevskaya, N., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S. & Dumas, P. (2000). The crystal structure of UUCG tetraloop. J. Mol. Biol. 304, 35–42.
- Famulare, M., Chang, S., Iber, J., Zhao, K., Adeniji, J.A., Bukbuk, D., Baba, M., Behrend, M., Burns, C.C.
  & Oberste, M.S. (2016). Sabin Vaccine Reversion in the Field: a Comprehensive Analysis of Sabin-Like Poliovirus Isolates in Nigeria. J. Virol. 90, 317–331.
- Ferner, J., Villa, A., Duchardt, E., Widjajakusuma, E., Wöhnert, J., Stock, G. & Schwalbe, H. (2008). NMR and MD studies of the temperature-dependent dynamics of RNA YNMG-tetraloops. *Nucleic Acids Res.* **36**, 1928–1940.
- Finger, L.D., Trantirek, L., Johansson, C. & Feigon, J. (2003). Solution structures of stem-loop RNAs that bind to the two N-terminal RNA-binding domains of nucleolin. *Nucleic Acids Res.* **31**, 6461–6472.
- Flanegan, J. & Van Dyke, T. (1979). Isolation of a soluble and template-dependent poliovirus RNA polymerase that copies virion RNA in vitro. *J Virol* **32**, 24212–24219.
- Flanegan, J.B., Petterson, R.F., Ambros, V., Hewlett, N.J. & Baltimore, D. (1977). Covalent linkage of a protein to a defined nucleotide sequence at the 5'-terminus of virion and replicative intermediate RNAs of poliovirus. *Proc. Natl. Acad. Sci. U. S. A.* 74, 961–5.

- Gamarnik, A. & Andino, R. (1997). Two functional complexes fromed by KH domain containing protens with the 5' noncoding region of poliovirus RNA. *RNA* **3**, 882–892.
- Gamarnik, A. V & Andino, R. (1998). Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev.* 12, 2293–2304.
- Gamarnik, A. V & Andino, R. (2000). Interactions of viral protein 3CD and poly(rC) binding protein with the 5' untranslated region of the poliovirus genome. *J Virol.* 74, 2219–2226.
- Giachetti, C., Hwang, S.S. & Semler, B.L. (1992). cis-acting lesions targeted to the hydrophobic domain of a poliovirus membrane protein involved in RNA replication. *J. Virol.* **66**, 6045–6057.
- Gohara, D.W., Crotty, S., Arnold, J.J., Yoder, J.D., Andino, R. & Cameron, C.E. (2000). Poliovirus RNAdependent RNA polymerase (3Dpol): structural, biochemical, and biological analysis of conserved structural motifs A and B. *J Biol Chem.* **275**, 25523–25532.
- Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J.W., Barclay, W. & Evans, D.J. (2000). Identification of a cis-acting replication element within the poliovirus coding region. J. Virol. 74, 4590–4600.
- Goodfellow, I., Polacek, C., Andino, R. & Evans, D. (2003a). The poliovirus 2C cis-acting replication element-mediated uridylylation of VPg is not required for synthesis of negative-sense genomes. *J. Gen. Virol.* **84**, 2359–63.
- Goodfellow, I.G., Kerrigan, D. & Evans, D.J. (2003b). Structure and function analysis of the poliovirus cisacting replication element (CRE). *RNA* 9, 124–137.
- Gradi, A., Svitkin, Y. V, Imataka, H. & Sonenberg, N. (1998). Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 11089–11094.
- Hämmerle, T., Hellen, C.U. & Wimmer, E. (1991). Site-directed mutagenesis of the putative catalytic triad of poliovirus 3C proteinase. J. Biol. Chem. 266, 5412–5416.
- Hämmerle, T., Molla, A. & Wimmer, E. (1992). Mutational analysis of the proposed FG loop of poliovirus proteinase 3C identifies amino acids that are necessary for 3CD cleavage and might be determinants of a function distinct from proteolytic activity. *J Virol.* **66**, 6028–6034.
- Han, J.-Q., Townsend, H.L., Jha, B.K., Paranjape, J.M., Silverman, R.H. & Barton, D.J. (2007). A phylogenetically conserved RNA structure in the poliovirus open reading frame inhibits the antiviral endoribonuclease RNase L. *J Virol.* **81**, 5561–5572.
- Hanecak, R., Semler, B.L., Anderson, C.W. & Wimmer, E. (1982). Proteolytic processing of poliovirus polypeptides: antibodies to polypeptide P3-7c inhibit cleavage at glutamine-glycine pairs. *Proc. Natl. Acad. Sci. U. S. A.* 79, 3973–3977.
- Hanecak, R., Semler, B.L., Ariga, H., Anderson, C.W. & Wimmer, E. (1984). Expression of a cloned gene segment of poliovirus in E. coli: evidence for autocatalytic production of the viral proteinase. *Cell* 37, 1063–1073.
- Hansen, J.L., Long, A.M. & Schultz, S.C. (1997). Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* 5, 1109–1122.
- Harber, J.J., Bradley, J., Anderson, C.W. & Wimmer, E. (1991). Catalysis of poliovirus VP0 maturation cleavage is not mediated by serine 10 of VP2. *J. Virol.* **65**, 326–334.
- Harris, K.S., Reddigari, S.R., Nicklin, M.J., Hämmerle, T. & Wimmer, E. (1992). Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. *J Virol.* **66**, 7481–7489.
- Harris, K.S., Xiang, W., Alexanders, L., Lane, W.S., Paul, A. V & Wimmer, E. (1994). Interaction of Poliovirus Polypeptide 3CDPro with the 5' and 3' Termini of the Poliovirus Genome. Identification of viral and celllular cofactors needed for efficient binding. *J Biol Chem* **269**, 27004–27014.
- He, Y., Bowman, V.D., Mueller, S., Bator, C.M., Bella, J., Peng, X., Baker, T.S., Wimmer, E., Kuhn, R.J. & Rossmann, M.G. (2000). Interaction of the poliovirus receptor with poliovirus. *Proc. Natl. Acad.*

Sci. U. S. A. 97, 79–84.

- He, Y., Mueller, S., Chipman, P.R., Bator, C.M., Peng, X., Bowman, V.D., Mukhopadhyay, S., Wimmer, E., Kuhn, R.J. & Rossmann, M.G. (2003). Complexes of poliovirus serotypes with their common cellular receptor, CD155. J. Virol. 77, 4827–4835.
- Headey, S.J., Huang, H., Claridge, J.K., Soares, G.A., Dutta, K., Schwalbe, M., Yang, D. & Pascal, S.M. (2007). NMR structure of stem-loop D from human rhinovirus-14. *RNA* **13**, 351–360.
- Hellen, C.U., Witherell, G.W., Schmid, M., Shin, S.H., Pestova, T. V, Gil, A. & Wimmer, E. (1993). A cytoplasmic 57-kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 7642–7646.
- Herold, J. & Andino, R. (2000). Poliovirus requires a precise 5' end for efficient positive-strand RNA synthesis. J. Virol. 74, 6394–6400.
- Herold, J. & Andino, R. (2001). Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol. Cell* 7, 581–591.
- Heus, H.A. & Pardi, A. (1991). Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science (80-. ).* **253**, 191–194.
- Hogle, J.M., Chow, M. & Filman, D.J. (1985). Three-dimensional structure of poliovirus at 2.9 A resolution. *Science* **229**, 1358–1365.
- Hope, D.A., Diamond, S.E. & Kirkegaard, K. (1997). Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB. *J. Virol.* **71**, 9490–9598.
- Humphrey, W., Dalke, A. and Schulten, K. (1996). VMD Visual Molecular Dynamics. J. Molec. Graph. 14, 33–38.
- Ihle, Y., Ohlenschläger, O., Häfner, S., Duchardt, E., Zacharias, M., Seitz, S., Zell, R., Ramachandran, R. & Görlach, M. (2005). A novel cGUUAg tetraloop structure with a conserved yYNMGg-type backbone conformation from cloverleaf 1 of bovine enterovirus 1 RNA. *Nucleic Acids Res.* 33, 2003–2011.
- Ivanoff, L. a, Towatari, T., Ray, J., Korant, B.D. & Petteway, S.R. (1986). Expression and site-specific mutagenesis of the poliovirus 3C protease in Escherichia coli. *Proc Natl Acad Sci U S A*. 83, 5392– 5396.
- Joachims, M., Van Breugel, P.C. & Lloyd, R.E. (1999). Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J. Virol.* **73**, 718–727.
- Johnson, K.L. & Sarnow, P. (1991). Three poliovirus 2B mutants exhibit noncomplementable defects in viral RNA amplification and display dosage-dependent dominance over wild-type poliovirus. J. Virol. 65, 4341–4349.
- Jones, S., Daley, D.T.A., Luscombe, N.M., Berman, H.M. & Thornton, J.M. (2001). Protein RNA interactions : a structural analysis. *Biochemistry* **29**, 943–954.
- Jore, J., De Geus, B., Jackson, R.J., Pouwels, P.H. & Enger-Valk, B.E. (1988). Poliovirus Protein 3CD Is the Active Protease for Processing of the Precursor Protein P1 in vitro. *J. Gen. Virol.* **69**, 1627–1636.
- Jucker, F.M. & Pardi, A. (1995). Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif. *Biochemistry* **34**, 14416–14427.
- Kean, K.M., Agut, H., Fichot, O., Wimmer, E. & Girard, M. (1988). A poliovirus mutant defective for selfcleavage at the COOH-terminus of the 3C protease exhibits secondary processing defects. *Virology* 163, 330–340.
- Keating, K.S., Toor, N. & Pyle, A.M. (2008). The GANC tetraloop: a novel motif in the group IIC intron structure. *J. Mol. Biol.* **383**, 475–481.
- Kempf, B.J. & Barton, D.J. (2008a). Poly(rC) binding proteins and the 5' cloverleaf of uncapped poliovirus mRNA function during de novo assembly of polysomes. *J. Virol.* **82**, 5835–5846.
- Kempf, B.J. & Barton, D.J. (2008b). Poliovirus 2A Pro Increases Viral mRNA and Polysome Stability

Coordinately in Time with Cleavage of eIF4G. J. Virol. 82, 5847–5859.

- Kempf, B.J., Kelly, M.M., Springer, C.L., Peersen, O.B. & Barton, D.J. (2013). Structural features of a picornavirus polymerase involved in the polyadenylation of viral RNA. *J. Virol.* **87**, 5629–5644.
- Kuyumcu-Martinez, N.M., Van Eden, M.E., Younan, P. & Lloyd, R.E. (2004). Cleavage of poly(A)-binding protein by poliovirus 3C protease inhibits host cell translation: a novel mechanism for host translation shutoff. *Mol. Cell. Biol.* **24**, 1779–1790.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–5.
- Lama, J., Paul, a V, Harris, K.S. & Wimmer, E. (1994). Properties of purified recombinant poliovirus protein 3AB as substrate for viral proteinases and as co-factor for RNA polymerase 3Dpol. *J. Biol. Chem.* **269**, 66–70.
- Lama, J., Sanz, M.A. & Rodríguez, P.L. (1995). A role for 3AB protein in poliovirus genome replication. *J. Biol. Chem.* **270**, 14430–14438.
- Langereis, M.A., Feng, Q., Nelissen, F.H.T., Virgen-Slane, R., van der Heden van Noort, G.J., Maciejewski, S., Filippov, D. V, Semler, B.L., van Delft, F.L. & van Kuppeveld, F.J.M. (2014). Modification of picornavirus genomic RNA using "click" chemistry shows that unlinking of the VPg peptide is dispensable for translation and replication of the incoming viral RNA. *Nucleic Acids Res.* 42, 2473– 2482.
- Lebars, I., Lamontagne, B., Yoshizawa, S., Aboul-Elela, S. & Fourmy, D. (2001). Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing. *EMBO J.* 20, 7250–7258.
- Lee, Y.F., Nomoto, A., Detjen, B.M. & Wimmer, E. (1977). A protein covalently linked to poliovirus genome RNA. *Proc. Natl. Acad. Sci. U. S. A.* 74, 59–63.
- Leong, L.E.C., Walker, P.A. & Porter, A.G. (1993). Human Rhinovirus-14 Protease 3C (3CPr0) Binds Specifically to the 5'-Noncoding Region of the Viral RNA. J. Biol. Chem. 268, 25735–25739.
- Li, J. & Baltimore, D. (1988). Isolation of poliovirus 2C mutants defective in viral RNA synthesis. *J. Virol.* **62**, 4016–4021.
- Li, J.P. & Baltimore, D. (1990). An Intragenic Revertant of a Poliovirus 2C Mutant Has an Uncoating Defect **64**, 1102–1107.
- Liu, P., Li, L., Keane, S.C., Yang, D., Leibowitz, J.L. & Giedroc, D.P. (2009). Mouse hepatitis virus stemloop 2 adopts a uYNMG(U)a-like tetraloop structure that is highly functionally tolerant of base substitutions. J. Virol. 83, 12084–12093.
- Liu, Y., Franco, D., Paul, A. V & Wimmer, E. (2007). Tyrosine 3 of poliovirus terminal peptide VPg(3B) has an essential function in RNA replication in the context of its precursor protein, 3AB. *J. Virol.* **81**, 5669–5684.
- Liu, Y., Wang, C., Mueller, S., Paul, A. V, Wimmer, E. & Jiang, P. (2010). Direct interaction between two viral proteins, the nonstructural protein 2C and the capsid protein VP3, is required for enterovirus morphogenesis. *PLoS Pathog.* **6**, e1001066.
- Mach, O., Verma, H., Khandait, D.W., Sutter, R.W., O'Connor, P.M., Pallansch, M.A., Cochi, S.L., Linkins, R.W., Chu, S.Y., Wolff, C. & Jafari, H.S. (2014). Prevalence of asymptomatic poliovirus infection in older children and adults in northern India: analysis of contact and enhanced community surveillance, 2009. J. Infect. Dis. 210 Suppl, S252-258.
- Marcotte, L.L., Wass, A.B., Gohara, D.W., Pathak, H.B., Arnold, J.J., Filman, D.J., Cameron, C.E. & Hogle, J.M. (2007). Crystal structure of poliovirus 3CD protein: virally encoded protease and precursor to the RNA-dependent RNA polymerase. J. Virol. 81, 3583–3596.
- Melchers, W.J.G., Zoll, J., Tessari, M., Bakhmutov, D. V, Gmyl, A.P., Agol, V.I. & Heus, H. a. (2006). A GCUA tetranucleotide loop found in the poliovirus oriL by in vivo SELEX (un)expectedly forms a YNMG-like structure: Extending the YNMG family with GYYA. *RNA* **12**, 1671–1682.
- Mendelsohn, C.L., Wimmer, E. & Racaniello, V.R. (1989). Cellular receptor for poliovirus: molecular
cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. *Cell* **56**, 855–865.

- Miller, S.B., Yildiz, F.Z., Lo, J.A., Wang, B. & D'Souza, V.M. (2014). A structure-based mechanism for tRNA and retroviral RNA remodelling during primer annealing. *Nature* **515**, 591–595.
- Mohammed, S., Phelan, M.M., Rasul, U. & Ramesh, V. (2014). NMR elucidation of the role of Mg2+ in the structure and stability of the conserved RNA motifs of the EMCV IRES element. *Org. Biomol. Chem.* **12**, 1495–1509.
- Mohan, S., Hsiao, C., Bowman, J.C., Wartell, R. & Williams, L.D. (2010). RNA tetraloop folding reveals tension between backbone restraints and molecular interactions. J. Am. Chem. Soc. 132, 12679–12689.
- Molla, A., Harris, K.S., Paul, A. V., Shin, S.H., Mugavero, J. & Wimmer, E. (1994). Stimulation of poliovirus proteinase 3Cpro-related proteolysis by the genome-linked protein VPg and its precursor 3AB. J. Biol. Chem. 269, 27015–27020.
- Montmayeur, A.M., Ng, T.F.F., Schmidt, A., Zhao, K., Magaña, L., Iber, J., Castro, C.J., Chen, Q., Henderson, E., Ramos, E., Shaw, J., Tatusov, R.L., Dybdahl-Sissoko, N., Endegue-Zanga, M.C., Adeniji, J.A., Oberste, M.S. & Burns, C.C. (2017). High-Throughput Next-Generation Sequencing of Polioviruses. J. Clin. Microbiol. 55, 606–615.
- Morasco, B.J., Sharma, N., Parilla, J. & Flanegan, J.B. (2003). Poliovirus cre(2C)-dependent synthesis of VPgpUpU is required for positive-but not negative-strand RNA synthesis. J. Virol. 77, 5136–5144.
- Mosimann, S.C., Cherney, M.M., Sia, S., Plotch, S. & James, M.N. (1997). Refined X-ray crystallographic structure of the poliovirus 3C gene product. J Mol Biol 273, 1032–1047.
- Moustafa, I.M., Gohara, D.W., Uchida, A., Yennawar, N. & Cameron, C.E. (2015). Conformational Ensemble of the Poliovirus 3CD Precursor Observed by MD Simulations and Confirmed by SAXS: A Strategy to Expand the Viral Proteome? *Viruses* 7, 5962–5986.
- Murray, K.E. & Barton, D.J. (2003). Poliovirus CRE-dependent VPg uridylylation is required for positivestrand RNA synthesis but not for negative-strand RNA synthesis. J. Virol. 77, 4739–4750.
- Murray, K.E., Roberts, A.W. & Barton, D.J. (2001). Poly(rC) binding proteins mediate poliovirus mRNA stability. *RNA* 7, 1126–1141.
- Muslin, C., Joffret, M.-L., Pelletier, I., Blondel, B. & Delpeyroux, F. (2015). Evolution and Emergence of Enteroviruses through Intra- and Inter-species Recombination: Plasticity and Phenotypic Impact of Modular Genetic Exchanges in the 5' Untranslated Region. *PLOS Pathog.* 11, e1005266.
- Nathanson, N. (2008). The pathogenesis of poliomyelitis: what we don't know. Adv. Virus Res. 71, 1-50.
- Neufeld, K., Richards, O. & Ehrenfeld, E. (1991). Purification, characterization, and comparison of poliovirus RNA polymerase from native and recombinant sources. *J Biol Chem* **266**, 24212–24219.
- Novak, J.E. & Kirkegaard, K. (1994). Coupling between genome translation and replication in an RNA virus. *Genes Dev.* **8**, 1726–1737.
- Novoa, I. & Carrasco, L. (1999). Cleavage of eukaryotic translation initiation factor 4G by exogenously added hybrid proteins containing poliovirus 2Apro in HeLa cells: effects on gene expression. *Mol. Cell. Biol.* **19**, 2445–2454.
- Nugent, C.I., Johnson, K.L., Sarnow, P. & Kirkegaard, K. (1999). Functional coupling between replication and packaging of poliovirus replicon RNA. J. Virol. **73**, 427–435.
- Oberstrass, F.C., Lee, A., Stefl, R., Janis, M., Chanfreau, G. & Allain, F.H.-T. (2006). Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. *Nat. Struct. Mol. Biol.* **13**, 160–167.
- Ogram, S.A., Spear, A., Sharma, N. & Flanegan, J.B. (2010). The 5'CL-PCBP RNP complex, 3' poly(A) tail and 2A(pro) are required for optimal translation of poliovirus RNA. *Virology* **397**, 14–22.
- Oh, H.S., Pathak, H.B., Goodfellow, I.G., Arnold, J.J. & Cameron, C.E. (2009). Insight into poliovirus genome replication and encapsidation obtained from studies of 3B-3C cleavage site mutants. J. Virol. 83, 9370–9387.
- Ohlenschläger, O., Wöhnert, J., Bucci, E., Seitz, S., Häfner, S., Ramachandran, R., Zell, R. & Görlach, M.

(2004). The structure of the stemloop D subdomain of coxsackievirus B3 cloverleaf RNA and its interaction with the proteinase 3C. *Structure* **12**, 237–248.

- van Ooij, M.J.M., Polacek, C., Glaudemans, D.H.R.F., Kuijpers, J., van Kuppeveld, F.J.M., Andino, R., Agol, V.I. & Melchers, W.J.G. (2006). Polyadenylation of genomic RNA and initiation of antigenomic RNA in a positive-strand RNA virus are controlled by the same cis-element. *Nucleic Acids Res.* 34, 2953–2965.
- Ostareck-Lederer, A., Ostareck, D.H. & Hentze, M.W. (1998). Cytoplasmic regulatory functions of the KHdomain proteins hnRNPs K and E1/E2. *Trends Biochem. Sci.* 23, 409–411.
- Pallansch, M.A., Kew, O.M., Semler, B.L., Omilianowski, D.R., Anderson, C.W., Wimmer, E. & Rueckert, R.R. (1984). Protein processing map of poliovirus. J. Virol. 49, 873–880.
- Parsley, T.B., Cornell, C.T. & Semler, B.L. (1999). Modulation of the RNA binding and protein processing activities of poliovirus polypeptide 3CD by the viral RNA polymerase domain. *J. Biol. Chem.* **274**, 12867–12876.
- Parsley, T.B., Towner, J.S., Blyn, L.B., Ehrenfeld, E. & Semler, B.L. (1997). Poly (rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA* **3**, 1124–1134.
- Paul, A. V, van Boom, H.J., Filippov, D. & Wimmer, E. (1998). Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature* 393, 280–284.
- Paul, A. V, Rieder, E., Kim, D.W., van Boom, H.J. & Wimmer, E. (2000). Identification of an RNA Hairpin in Poliovirus RNA That Serves as the Primary Template in the In Vitro Uridylylation of VPg. J. Virol. 74, 10359–10370.
- Pelletier, J., Kaplan, G., Racaniello, V.R. & Sonenberg, N. (1988). Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region. *Mol. Cell. Biol.* 8, 1103–1112.
- Pelletier, J. & Sonenberg, N. (1989). Internal Binding of Eucaryotic Ribosomes on Poliovirus RNA: Translation in HeLa Cell Extracts ESP-ZsE. J. Virol. 63, 441–444.
- Perera, R., Daijogo, S., Walter, B.L., Nguyen, J.H.C. & Semler, B.L. (2007). Cellular protein modification by poliovirus: the two faces of poly(rC)-binding protein. *J Virol* **81**, 8919–8932.
- Petrov, A.I., Zirbel, C.L. & Leontis, N.B. (2013). Automated classification of RNA 3D motifs and the RNA 3D Motif Atlas. *RNA* **19**, 1327–1340.
- Pfister, T. & Wimmer, E. (1999). Characterization of the nucleoside triphosphatase activity of poliovirus protein 2C reveals a mechanism by which guanidine inhibits poliovirus replication. *J. Biol. Chem.* **274**, 6992–7001.
- Pilipenko, E. V, Gmyl, A.P., Maslova, S. V, Svitkin, Y. V, Sinyakov, A.N. & Agol, V.I. (1992a). Prokaryotic-like cis elements in the cap-independent internal initiation of translation on picornavirus RNA. *Cell* **68**, 119–131.
- Pilipenko, E. V, Maslova, S. V, Sinyakov, A.N. & Agol, V.I. (1992b). Towards identification of cis-acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures. *Nucleic Acids Res.* **20**, 1739–1745.
- Pilipenko, E. V, Poperechny, K. V, Maslova, S. V, Melchers, W.J.G., Slot, H.J.B. & Agoll, V.I. (1996). Cis-element, oriR, involved in the initiation of (-) strand poliovirus RNA: a quasi-globular multidomain RNA structure maintained by tertiary ('kissing') interactions. *EMBO J.* 15, 5428–5436.
- Plotch, S.J. & Palant, O. (1995). Poliovirus protein 3AB forms a complex with and stimulates the activity of the viral RNA polymerase, 3Dpol. *J. Virol.* **69**, 7169–7179.
- Porter, D.C., Ansardi, D.C., Lentz, M.R. & Morrow, C.D. (1993). Expression of poliovirus P3 proteins using a recombinant vaccinia virus results in proteolytically active 3CD precursor protein without further processing to 3Cpro and 3Dpol. *Virus Res.* **29**, 241–254.
- Proctor, D.J., Schaak, J.E., Bevilacqua, J.M., Falzone, C.J. & Bevilacqua, P.C. (2002). Isolation and

characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions. *Biochemistry* **41**, 12062–12075.

- Prostova, M.A., Gmyl, A.P., Bakhmutov, D. V, Shishova, A.A., Khitrina, E. V, Kolesnikova, M.S., Serebryakova, M. V, Isaeva, O. V & Agol, V.I. (2015). Mutational robustness and resilience of a replicative cis-element of RNA virus: promiscuity, limitations, relevance. *RNA Biol.* 12, 1338–1354.
- Racaniello, V.R. (1996). Early events in poliovirus infection: virus-receptor interactions. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 11378–11381.
- Rieder, E., Paul, A. V., Dong Wook, K., van Boom, H.J. & Wimmer, E. (2000). Genetic and Biochemical Studies of Poliovirus cis-Acting Replication Element cre in Relation to VPg Uridylylation. J. Virol. 74, 10371–10380.
- Rieder, E., Xiang, W., Paul, A. & Wimmer, E. (2003). Analysis of the cloverleaf element in a human rhinovirus type 14/poliovirus chimera: correlation of subdomain D structure, ternary protein complex formation and virus replication. *J Gen Virol* **84**, 2203–2216.
- Rodríguez, P.L. & Carrasco, L. (1993). Poliovirus protein 2C has ATPase and GTPase activities. J. Biol. Chem. 268, 8105–10.
- Rodríguez, P.L. & Carrasco, L. (1995). Poliovirus protein 2C contains two regions involved in RNA binding activity. *J. Biol. Chem.* **270**, 10105–10112.
- Roehl, H.H., Parsley, T.B., Ho, T. V & Semler, B.L. (1997). Processing of a cellular polypeptide by 3CD proteinase is required for poliovirus ribonucleoprotein complex formation. *J. Virol.* **71**, 578–585.
- Rohll, J.B., Moon, D.H., Evans, D.J. & Almond, J.W. (1995). The 3' untranslated region of picornavirus RNA: features required for efficient genome replication. *J. Virol.* **69**, 7835–7844.
- SantaLucia, J., Kierzek, R. & Turner, D.H. (1992). Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin. *Science* **256**, 217–219.
- Schärpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S. & Rösch, P. (2000). Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. *Eur. J. Boichem.* 267, 2397–2408.
- Schulte, M.B., Draghi, J.A., Plotkin, J.B. & Andino, R. (2015). Experimentally guided models reveal replication principles that shape the mutation distribution of RNA viruses. *Elife* **4**.
- Schwalbe, M., Ohlenschläger, O., Marchanka, A., Ramachandran, R., Häfner, S., Heise, T. & Görlach, M. (2008). Solution structure of stem-loop alpha of the hepatitis B virus post-transcriptional regulatory element. *Nucleic Acids Res.* 36, 1681–1689.
- Sean, P., Nguyan, J.H.C. & Semler, B.L. (2008). The linker domain of poly(rC) binding protein 2 is a major determinant in poliovirus cap-independent translation. *Virology* 378, 243–253.
- Serra, M.J., Barnes, T.W., Betschart, K., Gutierrez, M.J., Sprouse, K.J., Riley, C.K., Stewart, L. & Temel, R.E. (1997). Improved Parameters for the Prediction of RNA Hairpin Stability † **2960**, 4844–4851.
- Sharma, N., O'Donnell, B.J., Flanegan, J.B., Donnell, B.J.O. & Flanegan, J.B. (2005). 3J-Terminal Sequence in Poliovirus Negative-Strand 3'-terminal sequence in poliovirus negative-strand templates is the primary cis-acting element required for VPgpUpU-primed positive-strand initiation. J. Virol. 79, 3565–3577.
- Sharma, N., Ogram, S.A., Morasco, B.J., Spear, A., Chapman, N.M. & Flanegan, J.B. (2009). Functional role of the 5' terminal cloverleaf in Coxsackievirus RNA replication. *Virology* **393**, 238–249.
- Sheehy, J.P., Davis, A.R. & Znosko, B.M. (2010). Thermodynamic characterization of naturally occurring RNA tetraloops. *RNA* 16, 417–429.
- Shen, M., Reitman, Z.J., Zhao, Y., Moustafa, I., Wang, Q., Arnold, J.J., Pathak, H.B. & Cameron, C.E. (2008). Picornavirus genome replication. Identification of the surface of the poliovirus (PV) 3C dimer that interacts with PV 3Dpol during VPg uridylylation and construction of a structural model for the PV 3C2-3Dpol complex. J. Biol. Chem. 283, 875–888.
- Shih, S., Chen, T. & Wu, C. (2004). Mutations at KFRDI and VGK Domains of Enterovirus 71 3C Protease

Affect Its RNA Binding and Proteolytic Activities. J. Biomed. Sci. 239–248.

- Shulman, L.M., Martin, J., Sofer, D., Burns, C.C., Manor, Y., Hindiyeh, M., Gavrilin, E., Wilton, T., Moran-Gilad, J., Gamzo, R., Mendelson, E., Grotto, I., Chen, Q., Dybdahl-Sissoko, N., Iber, J., Mandelbaum, M., Oberste, S., Penaranda, S., Rogers, S., Agabiev, I., Alfandari, J., Azar, R., Halmut, T., Indenbaum, V., Mandelbaum, M., Michaeli, M., Mor, O., Perepliotchikov, Y., Rom, D., Silberstein, I., Weil, M., Anis, E., Kaliner, E., Kopel, E., Singer-Shepherd, R., Dunn, G., Li, L. & Pfeifer, D. (2014). Genetic Analysis and Characterization of Wild Poliovirus Type 1 During Sustained Transmission in a Population With >95% Vaccine Coverage, Israel 2013. *Clin. Infect. Dis.* 60, 1057–64.
- Silvera, D., Gamarnik, a V & Andino, R. (1999). The N-terminal K homology domain of the poly(rC)binding protein is a major determinant for binding to the poliovirus 5'-untranslated region and acts as an inhibitor of viral translation. *J. Biol. Chem.* **274**, 38163–38170.
- Silvestri, L.S., Parilla, J.M., Morasco, B.J., Ogram, S.A. & Flanegan, J.B. (2006). Relationship between poliovirus negative-strand RNA synthesis and the length of the 3' poly(A) tail. *Virology* **345**, 509–519.
- Siomi, H., Choi, M., Siomi, M.C., Nussbaum, R.L. & Dreyfuss, G. (1994). Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. *Cell* 77, 33–39.
- Siomi, H., Matunis, M.J., Michael, W.M. & Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.* **21**, 1193–1198.
- Song, Y., Liu, Y., Ward, C.B., Mueller, S., Futcher, B., Skiena, S., Paul, A. V & Wimmer, E. (2012). Identification of two functionally redundant RNA elements in the coding sequence of poliovirus using computer-generated design. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14301–14307.
- Song, Y., Tzima, E., Ochs, K., Bassili, G., Trusheim, H., Linder, M., Preissner, K.T. & Niepmann, M. (2005). Evidence for an RNA chaperone function of polypyrimidine tract-binding protein in picornavirus translation. *RNA* 11, 1809–1824.
- Spear, A., Ogram, S.A., Morasco, B.J., Smerage, L.E. & Flanegan, J.B. (2015). Viral precursor protein P3 and its processed products perform discrete and essential functions in the poliovirus RNA replication complex. *Virology* 485, 491–501.
- Spector, D.H. & Baltimore, D. (1974). Requirement of 3'-terminal poly(adenylic acid) for the infectivity of poliovirus RNA. Proc. Natl. Acad. Sci. U. S. A. 71, 2983–2987.
- Steil, B.P., Kempf, B.J. & Barton, D.J. (2010). Poly(A) at the 3' end of positive-strand RNA and VPg-linked poly(U) at the 5' end of negative-strand RNA are reciprocal templates during replication of poliovirus RNA. J. Virol. 84, 2843–2858.
- Strauss, D.M. & Wuttke, D.S. (2007). Characterization of protein-protein interactions critical for poliovirus replication: analysis of 3AB and VPg binding to the RNA-dependent RNA polymerase. J. Virol. 81, 6369–6378.
- Svitkin, Y. V, Costa-Mattioli, M., Herdy, B., Perreault, S. & Sonenberg, N. (2007). Stimulation of picornavirus replication by the poly(A) tail in a cell-free extract is largely independent of the poly(A) binding protein (PABP). *RNA* 13, 2330–2340.
- Takeda, N., Yang, C.F., Kuhn, R.J. & Wimmer, E. (1987). Uridylylation of the genome-linked protein of poliovirus in vitro is dependent upon an endogenous RNA template. *Virus Res.* **8**, 193–204.
- Teterina, N.L., Gorbalenya, A.E., Egger, D., Bienz, K. & Ehrenfeld, E. (1997). Poliovirus 2C protein determinants of membrane binding and rearrangements in mammalian cells. *J. Virol.* **71**, 8962–8972.
- Teterina, N.L., Pinto, Y., Weaver, J.D., Jensen, K.S. & Ehrenfeld, E. (2011). Analysis of poliovirus protein 3A interactions with viral and cellular proteins in infected cells. *J. Virol.* **85**, 4284–4296.
- Theimer, C.A., Finger, L.D. & Feigon, J. (2003). YNMG tetraloop formation by a dyskeratosis congenita mutation in human telomerase RNA. *RNA* **9**, 1446–1455.
- Thompson, A.A. & Peersen, O.B. (2004). Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J.* 23, 3462–3471.

- Todd, S., Towner, J.S., Brown, D.M. & Semler, B.L. (1997). Replication-competent picornaviruses with complete genomic RNA 3' noncoding region deletions. *J. Virol.* **71**, 8868–8874.
- Tolskaya, E.A., Romanova, L.I., Kolesnikova, M.S., Gmyl, A.P., Gorbalenya, A.E. & Agol, V.I. (1994). Genetic studies on the poliovirus 2C protein, an NTPase. A plausible mechanism of guanidine effect on the 2C function and evidence for the importance of 2C oligomerization. *J. Mol. Biol.* **236**, 1310–1323.
- Towner, J.S., Mazanet, M.M. & Semler, B.L. (1998). Rescue of defective poliovirus RNA replication by 3AB-containing precursor polyproteins. *J. Virol.* **72**, 7191–200.
- Toyoda, H., Nicklin, M.J., Murray, M.G., Anderson, C.W., Dunn, J.J., Studier, F.W. & Wimmer, E. (1986). A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. *Cell* **45**, 761–770.
- Trono, D., Andino, R. & Baltimore, D. (1988). An RNA sequence of hundreds of nucleotides at the 5' end of poliovirus RNA is involved in allowing viral protein synthesis. *J. Virol.* **62**, 2291–2299.
- Troxler, M., Egger, D., Pfister, T. & Bienz, K. (1992). Intracellular localization of poliovirus RNA by in situ hybridization at the ultrastructural level using single-stranded riboprobes. *Virology* **191**, 687–97.
- Vance, L.M., Moscufo, N., Chow, M. & Heinz, B.A. (1997). Poliovirus 2C region functions during encapsidation of viral RNA. J. Virol. 71, 8759–8765.
- Varani, G. (1995). Exceptionally stable nucleic acid hairpins 379–404.
- Varani, G., Cheong, C. & Tinoco, I. (1991). Structure of an unusually stable RNA hairpin. *Biochemistry* **30**, 3280–3289.
- Ventoso, I. & Carrasco, L. (1995). A poliovirus 2A(pro) mutant unable to cleave 3CD shows inefficient viral protein synthesis and transactivation defects. J. Virol. 69, 6280–6288.
- Vogt, D.A. & Andino, R. (2010). An RNA element at the 5'-end of the poliovirus genome functions as a general promoter for RNA synthesis. *PLoS Pathog.* 6, e1000936.
- Walter, B.L., Parsley, T.B., Ehrenfeld, E. & Semler, B.L. (2002). Distinct poly(rC) binding protein KH domain determinants for poliovirus translation initiation and viral RNA replication. J. Virol. 76, 12008–12022.
- Wang, C., Jiang, P., Sand, C., Paul, A. V & Wimmer, E. (2012). Alanine scanning of poliovirus 2CATPase reveals new genetic evidence that capsid protein/2CATPase interactions are essential for morphogenesis. J. Virol. 86, 9964–9975.
- Wang, C., Ma, H.-C., Wimmer, E., Jiang, P. & Paul, A. V. (2014). A C-terminal, cysteine-rich site in poliovirus 2C(ATPase) is required for morphogenesis. J. Gen. Virol. 95, 1255–65.
- Wang, Z., Day, N., Trifillis, P. & Kiledjian, M. (1999). An mRNA stability complex functions with poly(A)binding protein to stabilize mRNA in vitro. *Mol. Cell. Biol.* 19, 4552–4560.
- Wang, Z., Hartman, E., Roy, K., Chanfreau, G. & Feigon, J. (2011). Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. *Structure* 19, 999–1010.
- Woese, C.R., Gutell, R., Gupta, R. & Noller, H.F. (1983). Detailed Analysis of the Higher-Order Structure of 16S-Like Ribosomal Ribonucleic Acids. *Structure* 47, 621–669.
- Woese, C.R., Winker, S., Gutell, R.R., Winkers, S. & Gutell, R.R. (1990). Architecture of ribosomal RNA : Constraints on the sequence of "tetra-loops." *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8467–8471.
- Wu, H., Henras, A., Chanfreau, G. & Feigon, J. (2004). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. Amino Acids 101, 8307–8312.
- Wu, H., Yang, P.K., Butcher, S.E., Kang, S., Chanfreau, G. & Feigon, J. (2001). A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. *EMBO J.* 20, 7240–9.
- Xiang, W., Cuconati, A., Hope, D., Kirkegaard, K. & Wimmer, E. (1998). Complete protein linkage map

of poliovirus P3 proteins: interaction of polymerase 3Dpol with VPg and with genetic variants of 3AB. *J. Virol.* **72**, 6732–6741.

- Xiang, W., Cuconati, A., Paul, A. V., Cao, X. & Wimmer, E. (1995a). Molecular dissection of the multifunctional poliovirus RNA-binding protein 3AB. *RNA* 1, 892–904.
- Xiang, W., Harris, K.S., Alexander, L. & Wimmer, E. (1995b). Interaction between the 5'-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication. *J. Virol.* **69**, 3658–3667.
- Yakovenko, M.L., Gmyl, A.P., Ivanova, O.E., Eremeeva, T.P., Ivanov, A.P., Prostova, M.A., Baykova, O.Y., Isaeva, O. V., Lipskaya, G.Y., Shakaryan, A.K., Kew, O.M., Deshpande, J.M. & Agol, V.I. (2014). The 2010 outbreak of poliomyelitis in Tajikistan: epidemiology and lessons learnt. *Euro Surveill. Eur. Commun. Dis. Bull.* 19, 20706.
- Yin, J., Paul, A. V, Wimmer, E. & Rieder, E. (2003). Functional dissection of a poliovirus cis-acting replication element [PV-cre(2C)]: analysis of single- and dual-cre viral genomes and proteins that bind specifically to PV-cre RNA. J. Virol. 77, 5152–5166.
- Yogo, Y. & Wimmer, E. (1972). Polyadenylic acid at the 3'-terminus of poliovirus RNA. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 1877–1882.
- Ypma-Wong, M.F., Dewalt, P.G., Johnson, V.H., Lamb, J.G. & Semler, B.L. (1988). Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. *Virology* 166, 265– 270.
- Ypma-Wong, M.F. & Semler, B.L. (1987). Processing determinants required for in vitro cleavage of the poliovirus P1 precursor to capsid proteins. J. Virol. 61, 3181–3189.
- Zell, R., Sidigi, K., Bucci, E., Stelzner, A. & Görlach, M. (2002). Determinants of the recognition of enteroviral cloverleaf RNA by coxsackievirus B3 proteinase 3C. *RNA* **8**, 188–201.
- Zhang, P., Mueller, S., Morais, M.C., Bator, C.M., Bowman, V.D., Hafenstein, S., Wimmer, E. & Rossmann, M.G. (2008). Crystal structure of CD155 and electron microscopic studies of its complexes with polioviruses. *Proc. Natl. Acad. Sci. U. S. A.* 105, 18284–18289.
- Zhao, Q., Huang, H.-C., Nagaswamy, U., Xia, Y., Gao, X. & Fox, G.E. (2012). UNAC tetraloops: to what extent do they mimic GNRA tetraloops? *Biopolymers* 97, 617–628.
- Zheng, D.P., Zhang, L.B., Fang, Z.Y., Yang, C.F., Mulders, M., Pallansch, M.A. & Kew, O.M. (1993). Distribution of wild type 1 poliovirus genotypes in China. J. Infect. Dis. 168, 1361–7.
- Zhou, J., Bean, R.L., Vogt, V.M. & Summers, M. (2007). Solution structure of the Rous sarcoma virus nucleocapsid protein: muPsi RNA packaging signal complex. *J Mol Biol.* **365**, 453–467.
- Zhou, J., McAllen, J.K., Tailor, Y. & Summers, M.F. (2005). High affinity nucleocapsid protein binding to the muPsi RNA packaging signal of Rous sarcoma virus. *J. Mol. Biol.* **349**, 976–988.