

ПЕТРОВА Дарья Петровна

БЕЛКИ ТРАНСПОРТА КРЕМНИЯ SiT

03.01.03 – молекулярная биология

Автореферат
диссертации на соискание учёной степени
кандидата биологических наук

Новосибирск – 2012

Работа выполнена в Лимнологическом институте СО РАН

Научный руководитель:

д.х.н., профессор, академик **Грачев Михаил Александрович**

Официальные оппоненты:

Нетёсов Сергей Викторович, д.б.н., профессор, чл.-корр. РАН
Новосибирский государственный университет, проректор по
научной работе

Бондарь Александр Анатольевич, к.х.н.
Институт химической биологии и фундаментальной медицины СО
РАН, старший научный сотрудник

Ведущая организация:

Сибирский институт физиологии и биохимии растений СО РАН

Защита состоится « 1 » июня 2012 г. в 10:00 часов
на заседании диссертационного совета Д 003.045.01 на базе
Института химической биологии и фундаментальной медицины СО РАН
по адресу: 630090, Новосибирск, проспект академика Лаврентьева, 8

С диссертацией можно ознакомиться в библиотеке
Института химической биологии и фундаментальной медицины СО РАН

Автореферат разослан « ____ » _____ 2012 г.

Учёный секретарь диссертационного совета
к.х.н., доцент



Коваль В. В.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. Диатомовые водоросли (Bacillariophyta) – одноклеточные фотосинтезирующие эукариотические организмы, по некоторым данным число их видов может достигать 10^5 - 10^6 (Mann, 1994; Round, 1996). В современных водоемах разного типа диатомеи являются одной из наиболее распространенных групп микроводорослей. Морские диатомеи играют важную роль в биосфере, обеспечивая до 20% органического вещества Земли (Nelson *et al.*, 1995), что сопоставимо с первичной продукцией тропических лесов (Field *et al.*, 1998).

Отличительной чертой диатомей является их клеточная стенка, построенная из аморфного кремнезема. Кремнистый панцирь диатомовых водорослей – один из примеров микро- и наноструктурированного природного материала. Виды диатомей классифицируются согласно симметрии панцирей и их тонкой структуре, представленной порами, выростами, шипами. Формирование кремнистых панцирей диатомей происходит внутри клетки в специализированной органелле – везикуле отложения кремнезема (Silica Deposition Vesicle - SDV) (Drum, Pankratz, 1964; Reimann, 1964). Механизмы, регулирующие у диатомей процессы захвата кремния из водной среды, его внутриклеточного транспорта к SDV и отложения биогенного кремнезема, остаются неизвестными. Современная молекулярная биология пока не может объяснить, как диатомеи преобразуют закодированную в геномах информацию в конструкцию кремнистых панцирей.

Важным достижением в молекулярной биологии диатомовых водорослей было открытие нового класса мембранных белков - транспортеров кремния (Silicon Transporters – SITs) (Hildebrand *et al.*, 1997)*.

Первая работа была выполнена на морском центрическом виде *Cylindrotheca fusiformis*. Следующей диатомеей, у которой был обнаружен ген *sit*, была пресноводная пеннатная диатомея *Synedra acus* (Грачев и др., 2002). В дальнейшем гены *sit* изучали параллельно в нескольких лабораториях мира. Однако до настоящего времени ни один белок SIT не был выделен в чистом виде и не подвергнут экспериментальным структурно-функциональным исследованиям.

* Имеются два подхода, относящиеся к участию элементов в природных процессах. В геохимии принято обозначать именно элемент, который участвует в том или ином геохимическом цикле, например, говорить о круговороте кремния. Форма, в которой элемент участвует в процессе, зачастую точно не известна. В настоящей работе мы говорим о транспорте кремния, понимая, что в действительности из окружающей среды в клетку поступает не кремний, а свободная кремниевая кислота ($\text{Si}(\text{OH})_4$), а створка диатомей построена не из кремния, а из кремнезема ($\text{SiO}_2 \cdot n\text{H}_2\text{O}$).

Цель и задачи исследования. Целью настоящей работы было исследование белков SIT.

Были поставлены следующие **задачи**:

1. Провести сравнительный анализ предсказанных аминокислотных последовательностей белков SIT.

2. Выявить консервативные мотивы и дать анализ имеющихся гипотез о связи структуры белков SIT с их функцией.

3. Обнаружить белок SIT в протеоме *S. acus* subsp. *radians* путем иммуноблоттинга и разработать метод его выделения с помощью иммунохроматографии.

4. Расшифровать последовательность нуклеотидов полноразмерного гена *sit* *S. acus* subsp. *radians* с использованием техники геномного пиросеквенирования.

Научная новизна. Методом иммуноблоттинга белок SIT впервые обнаружен в протеоме диатомовой водоросли. В геноме *S. acus* subsp. *radians* впервые определена полная нуклеотидная последовательность гена *sit2*, для этого была использована техника пиросеквенирования. Впервые проведен биоинформационный анализ широкого круга аминокислотных последовательностей белков SIT диатомовых и некоторых хризофитовых водорослей, выявлены их консервативные особенности, которые, по-видимому, имеют значение для осуществления их функций.

Теоретическая и практическая значимость работы. Результаты, полученные в работе, позволяют планировать ведение дальнейших исследований в двух направлениях:

1 – Эксперименты по химической модификации белков SIT в составе живой клетки, например, при воздействии алкилирующими агентами, про которые известно, что они ингибируют транспорт кремния и рост диатомей. Воспользовавшись этим методом, можно будет попытаться идентифицировать точку алкилирования, выделив для этих целей алкилированный белок в чистом виде; изучить действие бифункциональных реагентов для расшифровки взаимоположения элементов вторичной структуры белка SIT.

2 – Нарботать расшифрованный в работе полный ген *sit2* для использования в бесклеточной системе, что позволит препаративно синтезировать белок SIT2, встроенный в мембрану и находящийся в нативном состоянии. Имея такой препарат, будет возможно получить информацию о белке SIT с помощью метода ЯМР, а также предпринять попытку кристаллизации белка для его рентгено-структурного анализа. Подобные

возможности являются принципиально важными в связи с тем, что белки SIT не имеют близких аналогов среди белков живых систем.

Разработанный метод выделения белков из биомассы *S. acus* subsp. *radians* может быть рекомендован для исследования протеомов других диатомовых водорослей. Полученные антисыворотки против белка SIT были использованы в нашей совместной работе с Бедошвилли Е.Д. и Лихошвай Е.В. и позволили впервые обнаружить белок SIT на ультратонких срезах диатомеи *S. acus* subsp. *radians* методом иммуноэлектронной микроскопии.

Апробация работы. Материалы диссертации были представлены на Международном симпозиуме «Живая клетка диатомей» (Иркутск, 2004), Международном рабочем совещании по клеточной физиологии «Transport mechanisms across cell membranes: channels and pumps» (Санкт-Петербург, 2004), 2-ом Российском симпозиуме «Химия и биология пептидов» (Санкт-Петербург, 2005), 9-ой Школе диатомологов России и стран СНГ «Морфология, систематика, онтогенез, экология и биогеография диатомовых водорослей» (Борок, 2005), 1(9)-ой Международной конференции молодых ботаников (Санкт-Петербург, 2006), 19th International Diatom Symposium (Иркутск, 2006), 4-ом съезде Российского общества биохимиков и молекулярных биологов (Новосибирск, 2008) и 5-ой Верещагинской байкальской конференции (Иркутск, 2010).

Публикации. По материалам диссертационной работы опубликовано 14 научных работ, в том числе 6 статей в изданиях из Перечня ВАК РФ.

Структура и объём работы. Диссертация состоит из введения, трёх глав (обзора литературы, описания объектов и методов исследования, результатов исследования и обсуждения), заключения, выводов, списка литературы, который включает 131 источник, и 4 приложений. Работа изложена на 137 страницах, содержит 26 рисунков и 9 таблиц.

СОДЕРЖАНИЕ РАБОТЫ

Анализ аминокислотных последовательностей SIT, предсказанных по последовательностям нуклеотидов генов *sit*

На момент начала настоящей работы (2002 г.) в международной базе данных GenBank было опубликовано 8 генов и фрагментов генов *sit* четырех видов диатомей (*Cylindrotheca fusiformis*, *Synedra acus*, *Phaeodactylum tricornutum* и *Chaetoceros muelleri*).

Анализ выровненных аминокислотных последовательностей, предсказанных по последовательностям нуклеотидов этих генов, показал, что SIT-белки имеют сходства и различия по аминокислотным последовательностям. Консервативным мотивом оказалась последовательность

CMLD, находящаяся между 4 и 5 трансмембранными доменами и предложенная на роль одного из компонентов активного центра транспорта кремниевой кислоты (Грачев и др., 2002; Петрова, Грачев, Щербакова, 2004). Согласно высказанной гипотезе (Grachev *et al.*, 2005) мотив CMLD в белках SIT связывает ион цинка, который, в свою очередь, служит акцептором кремниевой кислоты.

В рамках настоящей работы были вычислены и проанализированы профили гидропатии указанных выше аминокислотных последовательностей. Оказалось, что, несмотря на большое число аминокислотных замен, профили гидропатии фрагментов белков весьма сходны, при этом во всех молекулах мотив CMLD был обнаружен в гомологичных положениях между двумя гидрофобными областями на коротком участке с малым индексом гидропатии.

Анализ распределения положительно заряженных остатков лизина (K) и аргинина (R) показал, что они сосредоточены в основном в гидрофильных областях, а трансмембранные домены, непосредственно примыкающие к мотиву CMLD, не содержат остатков K и R. Была обнаружена интересная особенность белков SIT. Оказалось, что мотив CMLD расположен в необычно длинных предсказанных триптических пептидах длиной от 50 до 110 а.о. Анализ встречаемости триптических пептидов различной длины в 400 известных белках диатомовых водорослей, депонированных в GenBank, показал, что лишь 59 из них могут давать столь протяженные трипсиновые пептиды (рис. 1). Протяженные участки последовательностей а.о., не содержащие положительно заряженных остатков K и R могут формировать пору для беспрепятственного транспорта кремниевой кислоты. Остатки лизина и аргинина из-за высокого сродства к кремнезему (Kröger *et al.*, 1999) могли бы блокировать транспорт кремниевой кислоты сквозь пору.

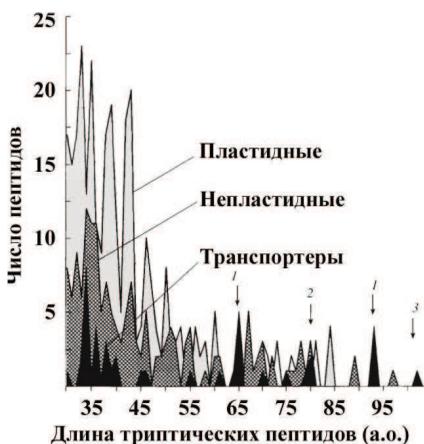


Рис. 1. Встречаемость трипсиновых пептидов для предсказанных по последовательностям нуклеотидов белков диатомовых водорослей. Вертикальными стрелками обозначены длинные трипсиновые пептиды белков SIT из *C. fusiformis* (1), *S. acus* (2) и *C. muelleri* (3).

В белках SIT, упомянутых выше четырех видов диатомей, был обнаружен еще один длинный триптический пептид, занимающий положение ближе к С-концу, который также может участвовать в формировании стенки транспортной поры.

В настоящее время (декабрь 2011 г.) в GenBank содержатся данные о фрагментах аминокислотных последовательностей 119 белков SIT из 58 видов диатомовых и 2 видов хризофитовых водорослей. Профили гидропатии белков хризофитовых и диатомовых водорослей оказались подобными (рис. 2), несмотря на значительное число аминокислотных замен. У белков SIT хризофитовых водорослей позиции 163-166 занимает мотив CMLD. Сходство профилей гидропатии, выявленное для столь широкого круга видов, очевидно, связано с тем, что все белки SIT имеют сходное расположение функционально значимых элементов в своих вторичных и третичных структурах, а наличие мотива CMLD у хризофитовых водорослей является дополнительным доводом в пользу его функциональной значимости.

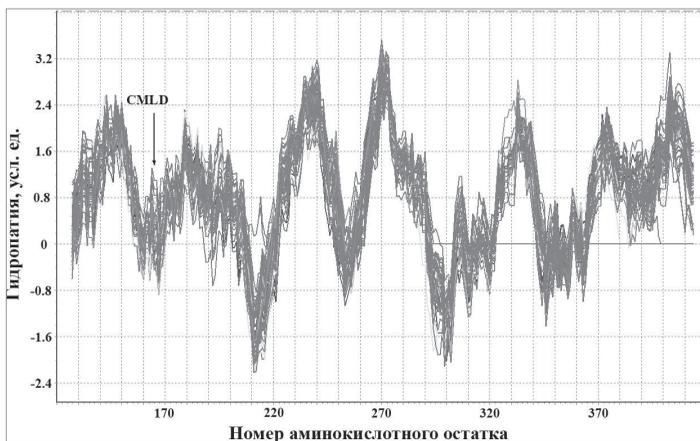


Рис. 2. Профили гидропатии гомологичных фрагментов предсказанных аминокислотных последовательностей 119 белков SIT из 58 видов диатомовых и 2 видов хризофитовых водорослей. Программа BioEdit. Нумерация аминокислотных остатков дана по последовательности CfSIT1. Цистеин мотива CMLD имеет порядковый номер 163.

Анализ встречаемости белков SIT, содержащих мотив CMLD, показал, что в своем каноническом виде он встречается в 39 предсказанных аминокислотных последовательности SIT у 15 видов диатомей и 2 видов хризофитовых водорослей. Отметим, что последовательность CMLD имеется без исключения во всех белках SIT пресноводных пеннатных диатомей, в том числе и у *S. acis* subsp. *radians*.

Его гомолог, в котором на месте остатка L (лейцин) присутствует остаток I (изолейцин), представлен в 45 последовательностях из 27 видов. В некоторых случаях вместо мотива CMLD присутствуют мотивы NMID, MMLD, MMID, QMID и SMID. Метионин (M), серин (S), глутамин (Q), гистидин (H) подобно цистеину, потенциально могут служить лигандами для иона цинка и благодаря этому участвовать в функции связывания кремниевой кислоты. В двух белках морской диатомеи *T. pseudonana* (TpSIT1 и TpSIT2) вместо CMLD присутствует мотив AMID (A – остаток аланина). Аланин не может служить лигандом для связывания иона цинка, поскольку в нем отсутствуют необходимые функциональные группы. Не исключено, что эти два белка выполняют другие функции.

В 2006 г. была высказана гипотеза (Thamatrakoln, Alverson, Hildebrand, 2006), что в состав активного центра транспортера кремния входят 4 мотива GxQ, которые связывают остаток кремниевой кислоты из окружающей водной среды. В рамках настоящей работы была проанализирована встречаемость мотива GxQ в 119 белках SIT. Сравнительный анализ показал, что наличие всех четырех мотивов GxQ не является консервативным признаком. Таким образом, эта гипотеза не выдержала проверки.

Рассмотренные выше последовательности аминокислот белков SIT являются предсказанными по нуклеотидным последовательностям генов. Наличие SIT в соответствующих протеомах не было показано.

Обнаружение белка SIT в протеоме *S. acus* subsp. *radians* и его выделение

Для того, чтобы выяснить связь структуры с функциями белков SIT и расшифровать молекулярные механизмы, обеспечивающие участие этих белков в транспорте кремния, необходимо провести комплекс исследований с помощью методов молекулярной биологии и биоорганической химии. Одной из задач настоящей работы была разработка метода, позволяющего извлечь из клеток *S. acus* subsp. *radians* ее как можно более полный протеом. Для исключения возможности потери белков за счет их сорбции на кремнеземе створок была применена техника растворения панцирей в 5 М NH₄F, pH 5,0. Как видно из **рисунка 3**, обработка диатомей указанным реагентом при 4°C в течение 12 ч приводит к полному растворению кремнистых створок, при этом лизис клеток не происходит.

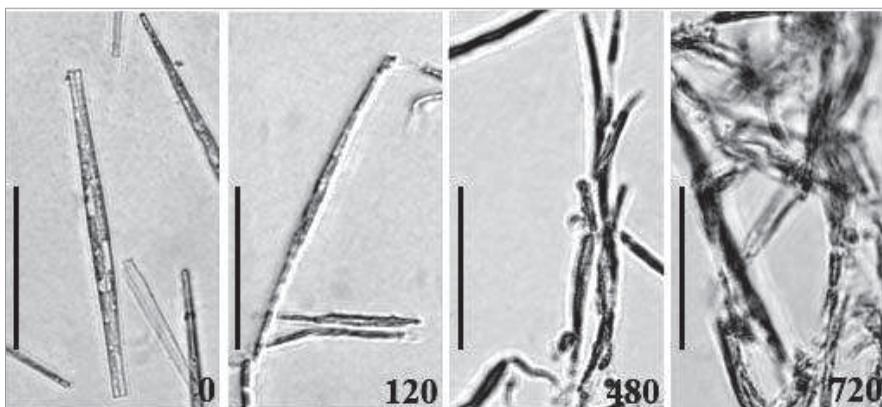


Рис. 3. Динамика растворения кремнистых панцирей диатомовой водоросли *S. acus* subsp. *radians* в 5 М NH_4F при pH 5,0 (световая микроскопия). Цифрами указано время в мин.

Лишенные кремнистых створок клетки далее экстрагировали буфером, содержащим 52 мМ β -меркаптоэтанол в 2%-ном растворе ДДС-Na / 250 мМ Трис-HCl, pH 8,8 при 37°C. В результате экстракции получен набор белков с широким диапазоном молекулярных масс - от 10 до 200 кДа. Стоит отметить, что при использовании коммерческого набора Total Protein Kit фирмы Bio-Rad для белковой экстракции на электрофореграмме выявлялся бедный спектр белков невысокой концентрации. Применение методов экстракции биомассы 2% ДДС-Na с нагреванием до 95°C также оказались малопродуктивными. Таким образом, разработанный метод оказался оптимальным для получения богатого набора белков протеома *S. acus* subsp. *radians*.

На основе данных о последовательности аминокислот белка SIT1 диатомеи *S. acus* subsp. *radians*, полученных ранее (Грачев и др., 2002) и в настоящей работе, руководствуясь рассчитанной вторичной структурой TmPred, были выбраны четыре потенциальные антигенные детерминанты, находящиеся во внемембранных участках (рис. 4). Для получения антисывороток против белков SIT В.В. Самуковым и др. (НПО «Вектор») были синтезированы пептиды P1-P4, мимикрирующие эти участки. Для обеспечения возможности конъюгации синтетических пептидов с альбумином в их С-концевые последовательности были введены остатки цистеина. В пептиде P3 на месте остатка цистеина (компонента мотива CMLD) был введен остаток серина (S), чтобы избежать осложнений, вызванных легкой окисляемостью цистеина. Пептиды были очищены методом ВЭЖХ. В результате иммунизации восьми кроликов конъюгатами пептидов с БСА были получены соответствующие

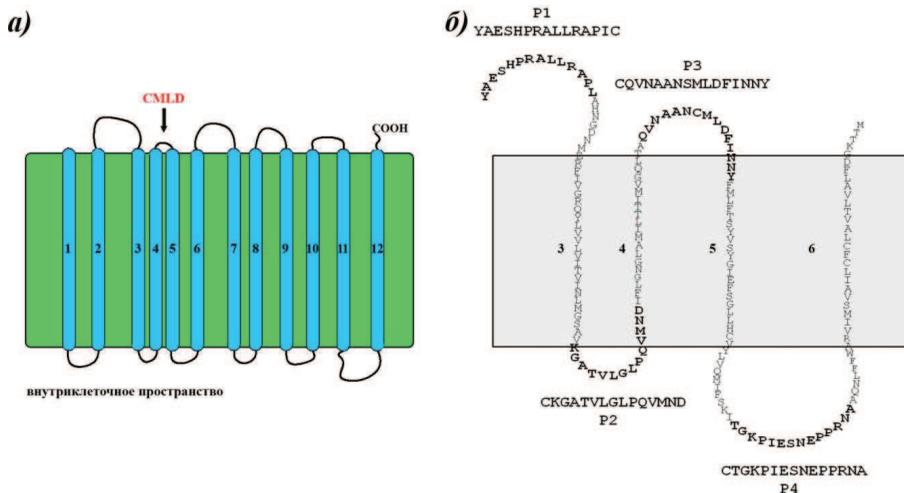


Рис. 4. Схема расположения в клеточной мембране белка SaSIT1: а) – расположение трансмембранных доменов и внемембранных участков по данным программы TMHMM; б) – потенциальные антигенные детерминанты выделены жирным шрифтом; P1-P4 – синтетические пептиды, использованные для получения антисывороток.

антисыворотки: anti-P1a, anti-P1b, anti-P2a, anti-P2b, anti-P3a, anti-P3b, anti-P4a и anti-P4b. Антисыворотки были истощены в отношении БСА.

Проверку специфичности связывания антисывороток с антигеном проводили дот-блот анализом, используя в качестве отрицательного контроля клеточный лизат *E. coli*. Следует отметить, что дот-блоттинг тормозился при преинкубации каждой из антисывороток с соответствующим синтетическим пептидом (P1-P4). По результатам раститровки, проведенной в дот-блот анализе, для дальнейшей работы были выбраны антисыворотки anti-P1a, anti-P1b, anti-P3a, anti-P3b, и anti-P4b, которые давали положительный ответ с антигеном в белковом экстракте, полученном после растворения кремнистого панциря из диатомей, и не показали неспецифической реакции с лизатом *E. coli*.

Анализ суммарного белкового экстракта *S. acus* subsp. *radians*, полученного в ходе работы, с использованием антисывороток anti-P1a, anti-P1b, anti-P3a, anti-P3b, и anti-P4b, а так же смесью антисывороток (anti-P1a, anti-P3a и anti-P4a), позволил выявить антиген с массой 62-64 кДа (**рис. 5**).

Для препаративного выделения белка SIT был применен метод иммунной хроматографии. Антисыворотку anti-P3a сорбировали на бромциан-сефарозу, далее, согласно инструкции производителя, проводили иммунохроматографию и в результате при промывке колонки буфером 100 мМ CH_3COONa / 500 мМ NaCl ,

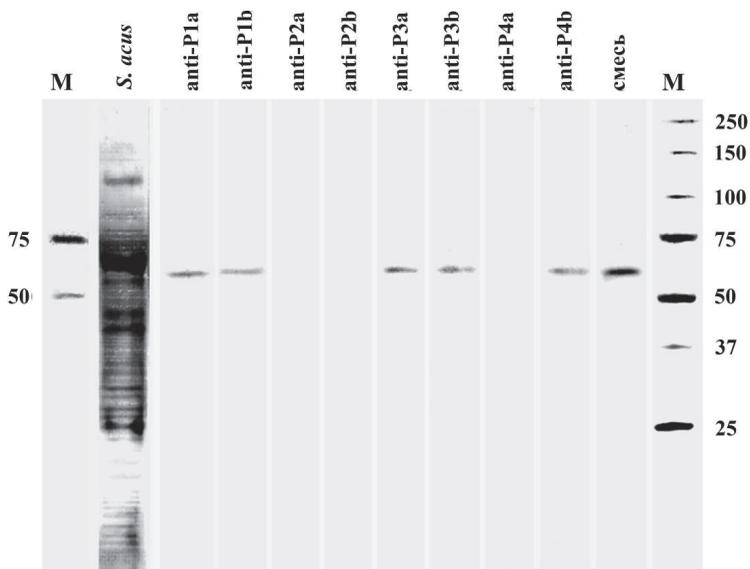


Рис. 5. Градиентный гель-электрофорез суммарной белковой фракции *S. acutus*, полученной после растворения кремнистого панциря (окраска Кумасси) и иммуноблоттинг с антителами против синтетических пептидов (смесь – anti-P1a, anti-P3a и anti-P4a; M – маркеры молекулярной массы).

pH 4,0, получили белок, который при гель-электрофорезе дал единственное детектируемое при окраске серебром пятно в районе молекулярных масс 62-64 кДа (рис. 6).

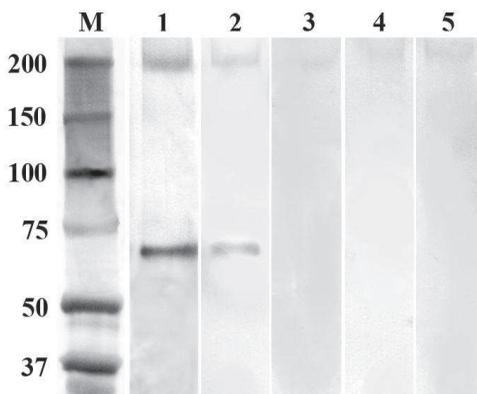


Рис. 6. Выделение белка SIT из суммарного белкового экстракта *S. acutus* subsp. *radians* методом иммунной хроматографии с антисывороткой anti-P3a. Градиентный гель-электрофорез, окраска серебром. M – маркер молекулярной массы; 1-5 - последовательные фракции, полученные в результате элюции колонки буфером 100 mM CH_3COONa / 500 mM NaCl , pH 4,0.

Выявление гена *sit2* в геноме *S. acus* subsp. *radians* и его анализ

Одной из задач настоящего исследования было получение полноразмерного гена *sit* *S. acus* subsp. *radians* для его последующего использования в молекулярно-биологических исследованиях. Расшифрованный ранее Т.А. Щербаковой и Ю.А. Масюковой (Лимнологический институт Сибирского отделения РАН) ген *sit* (FJ263063), который мы будем далее называть *Sasit1*, содержал на 5'-конце кодон GTT, кодирующий валин (V); стартовый кодон, кодирующий аминокислоту метионин (M) в нем найти не удалось. На 3'-конце присутствует стоп-кодон TAG. Имелись основания предполагать, что этот ген не является полноразмерным и может содержать нерасшифрованную нуклеотидную последовательность, расположенную «левее» упомянутого выше валинового кодона. Для поиска полноразмерного гена *Sasit1* была применена техника полногеномного пиросеквенирования. Высокомолекулярная ДНК была выделена из аксеничной культуры диатомеи *S. acus* subsp. *radians*. Библиотеку геномной ДНК для пиросеквенирования получали из 500 нг ДНК по стандартному протоколу фирмы Roche для секвенирования с реактивами серии Titanium (Rapid library preparation method manual, <http://www.my454.com>). Секвенирование проводили на пиросеквенаторе Roche Genome Sequencer GS FLX. Для сборки контигов были взяты последовательности, длина которых варьировала от 300 до 500 п.н. Сборку проводили с использованием ассемблера Newbler (Miller, Koren, Sutton, 2011) со стандартными настройками.

Пиросеквенирование (два раунда) и биоинформационный анализ полученных данных проведены в совместной работе с Ю.П. Галачьянцем и К.В. Хабудаевым (Лимнологический институт Сибирского отделения РАН). Контиг (37183) длиной 1069 п.н. имел высокое сходство с опубликованным ранее фрагментом последовательности гена, который дальше будем называть *Sasit1* (FJ263063). Предсказанная аминокислотная последовательность найденного фрагмента не содержала стартового метионина, имела по сравнению с фрагментом белка SaSIT1 единичную делецию и три точечные замены аминокислот (**рис. 7**).

Отличие вновь расшифрованного гена, который обозначен, как *Sasit1'*, может быть либо результатом неточностей секвенирования, либо следствием того, что ген *sit1'* является членом семейства генов, кодирующих белки SIT *S. acus* subsp. *radians*. Наряду с фрагментом гена *Sasit1'*, в одном из контигов (16677) длиной 2942 п.н. был обнаружен новый ген, сходный с геном *Sasit1*, который мы обозначили как *Sasit2*.

Поиск гена в контигах проводили по референсной последовательности *Sasit1*. Анализ предварительной сборки генома с помощью программного пакета Blast с параметрами blastn и e-value < 10⁻⁹ выявил сходную с *Sasit1* последовательность. Трансляция контига (16677) по всем рамкам считывания позволила найти непрерывную аминокислотную последовательность длиной 587 а.о. На расстоянии 17 а.о. с N-конца этой аминокислотной последовательности находится предположительно стартовый остаток метионина (M) белка SaSIT2. На 3'-конце гена был обнаружен терминирующий кодон TGA, а вслед за ним – некодирующий участок. Последним аминокислотным остатком белка SaSIT2 является глутаминовая кислота (E), кодируемая кодоном GAG. Общая длина белка SaSIT2 при такой интерпретации составляет 570 а.о. (рис. 7).

Аминокислотная последовательность, кодируемая геном *Sasit2*, содержит мотив CMLD в положении, гомологичном положению этого мотива в белках SaSIT1, SaSIT1' и других SIT - между двумя трансмембранными доменами. Протяженность не содержащего остатков R и K трипсинового пептида в SaSIT2 с мотивом CMLD составляет 75 а.о. Выявлены еще два трипсиновых пептида, имеющих длину 44 и 42 а.о. Таким образом, белок SaSIT2 сохраняет все предположительно важные для проявления его функций структурные особенности. Мотивы GxQ, предложенные Таматраколын и др. (Thamatrakoln, Alverson, Hildebrand, 2006) на роль участников активного центра транспорта кремния, в белке SaSIT2 найдены в четырех положениях.

Белки SaSIT1, SaSIT1' и SaSIT2 в гомологичных положениях содержат большое число (356) идентичных а.о. (рис. 7), на перекрывающихся участках аминокислотные последовательности оказались высоко сходными (84%), а.о. были идентичными в 71% положений. Эти последовательности демонстрируют сходство профилей гидропатии в перекрывающейся области. С-концевая часть белка SaSIT2 оказалась короче по сравнению с С-концевой последовательностью SaSIT1. Она не содержит coil-coiled мотивов. Число трансмембранных доменов SaSIT2, по расчетам программ SOSUI (Durbin *et al.*, 1998), TmPred (Hofmann, Stoffel, 1993), HMMTOP (Tusnády, Simon, 1998) и ТМНММ server v. 2.0 (Krogh *et al.*, 2001), составило 12 (табл. 1), мотив CMLD локализован между 6 и 7 трансмембранными доменами и находится вне клетки, что подтверждает его возможную значимость в связывании кремниевой кислоты из окружающей среды.

	5	15	25	35	45	55
SaSIT1	-----	-----	-----	-----	-----	-----
SaSIT1'	-----	-----	-----	-----	-----	-----
SaSIT2	MFIVARIASP	TYGKDDANIF	GISDGFQSF	NSGLTGAVIT	TVIGSLAWRI	IASSPFPLAFL

	65	75	85	95	105	115
SaSIT1	-----	VCLVLENVGI	CSGAWVLGRF	NKLIAGYQPD	EVYLEGAEKH	TSPEVTRRDG
SaSIT1'	-----	VCLVLENVGI	CSGAWVLGRF	NKLIAGYQPD	EVYLEGAERH	TSAPVTKRDK
SaSIT2	SNPLIYIIIR	VCLVLESVGI	CSGAWVLGRF	NKLIAGYQPD	EVYLEGAERH	TSAPVTKRDK

	125	135	145	155	165	175
SaSIT1	<=====	=====	=====	=====	=====	=====
SaSIT1'	DIDRLLTVFK	FIYSSALLVL	CLIIVHAAME	TTQTTSQADF	GLHPAVSFLI	FWFLIIWLAM
SaSIT2	DIDIAMTIVK	YLYSTGLLVF	SVVIVMSAMF	TQOTKIAND-	-AHPAVAFFL	FWFLIVWLAM

	185	195	205	215	225	235
SaSIT1	MEGGQGALVG	LQPVDKALYA	ESHPRALLRA	PLAHNGDNME	RFIVGRQYLV	VLVIIVINLM
SaSIT1'	MEGGQGALVG	LQPVDKALYA	ESHPRALLNT	KLAHNGDNME	RFIVGRQYLV	VLVIIVINLM
SaSIT2	MEGGQGCLIG	LKPIDKSLYK	DSHKVITYTNC	MLAHKGDNME	RFIVGRQFLV	VLVIIVINLC

	245	255	265	275	285	295
SaSIT1	====>	внутри	=====	====>	наружу	<=====
SaSIT1'	GSAYKGATVL	GLPQVMNDIF	LGNGLAMILT	TIMVGLTAQ	VNAAN	CMLD
SaSIT2	GSSLNKAKVL	GLDDTVEIF	LGSGVAMILT	TIILGQLTAQ	VNAAN	CMLD

	305	315	325	335	345	355
SaSIT1	YVSYGIEFSG	LLHCVYLVQM	IFSKITGKPI	ESNEPPRNAA	QNLFFWARVI	MSVAILCFCL
SaSIT1'	YVSYGIEFSG	LLHCVYLVQM	IFSKITGKPI	ESNEPPRNAA	QNLFFWARVI	MSVAILCFCL
SaSIT2	YVSLGIEFSG	LLHAVYLVQI	VFSKITGKPI	ESNEPPRSVA	QNLFFWARVF	VSLVILSFSL

	365	375	385	395	405	415
SaSIT1	====>	наружу	<=====	=====	внутри	=====
SaSIT1'	AVTLVALFDG	KT-MWDGVPP	YVSVIILFLL	MAVVGMMEGM	QIALFAVVMN	PEEELRNSPI
SaSIT2	AVTIVALFDR	KTSMYDGVPP	YVSVIVFFVL	MCVVGMMEGM	QIALFAVINL	PEDELKNYSM

	425	435	445	455	465	475
SaSIT1	AYANCQLTFA	GQNLQAFLLG	RQIFVATCMF	IVARIASPTY	GKDDANIFGI	SDGFQSFLLNS
SaSIT1'	AYANCQLTFA	GQNLQAFLLG	RQIFVATCMF	IVARIASPTY	GKDDANIFGI	SDGFQSFLLNS
SaSIT2	ASKNCTLAFQ	GTNLQSFLLG	RQIFVATCMF	VVARIASPSY	GSSDANIFGV	SDGFQTFLLNT

	485	495	505	515	525	535
SaSIT1	=====	внутри	=====	====>	наружу	<=
SaSIT1'	GLTGAVITTV	IGSLAWRIIA	SSFPLAFLSN	PLIYIIIRVC	LVLESVIGCS	GAWVFRFNK
SaSIT2	GLTGAVITTV	IGSLAWRIIA	SSFPLAFLSN	PLIYIIIRVC	LLLDAIGLCS	AAWLLALHK

	545	555	565	575	585	595
SaSIT1	LIAGYQPD	YLEGAERHTS	APVTKRDKDM	DIAMTTVKYL	YSTGLLVFSV	VIVMSAMFTQ
SaSIT1'	-----	-----	-----	-----	-----	-----
SaSIT2	QLVGFQID	YIG---KPSE	RAAAEKAVDE	ELAQE	-----	-----

	605	615	625	635	645	655
SaSIT1	QTEIANDAHP	AVALLFVLVS	FIVWARLIME	GGSGWLDRFL	NL	-----
SaSIT1'	-----	-----	-----	-----	-----	-----
SaSIT2	-----	-----	-----	-----	-----	-----

Рис. 7. Аминокислотные последовательности белков SaSIT1, SaSIT1' и SaSIT2 диатомовой водоросли *S. acus* subsp. *radians*. Серым цветом выделены консервативные участки, черным – мотив CMLD; === - трансмембранные домены.

Таблица 1. Структурные характеристики полных последовательностей белков SIT диатомовых водорослей.

Вид	Условное обозначение	Протяженность, а.о.	Масса, кДа	Coil-coiled домен	Число транс-мембранных доменов по TmPred	CMLD или его гомологи
<i>C. fusiformis</i>	CfSIT1	548	60,6	+	10	CMLD
	CfSIT2	543	59,5	+	10	CMLD
	CfSIT3	557	61	+	10	CMLD
	CfSIT4	548	60,5	+	10	CMLD
	CfSIT5	538	59,5	+	10	CMLD
<i>N. alba</i>	NaSIT	549	60,8	+	10	CMLD
<i>P. tricornutum</i>	PtSIT1	512	55,4	-	10	CMLD
	PtSIT2-1	506	54,9	-	10	CMLD
	PtSIT2-2	506	54,9	-	10	CMLD
	PtSIT3	599	65	-	10	CMLD
	PtSIT4	512	55,3	-	10	CMLD
<i>T. pseudonana</i>	TpSIT1	517	55,8	-	10	AMID
	TpSIT2	528	57,2	-	10	AMID
	TpSIT3	565	62,1	-	10	CMID
<i>S. costatum</i>	ScSIT	508	55,4	-	10	CMLD
<i>S. acus</i> subs. <i>radians</i>	SaSIT2	570	62,4	-	12	CMLD

К настоящему времени в GenBank опубликовано 15 аминокислотных последовательностей полноразмерных белков SIT из 6 видов диатомей (Hildebrand *et al.*, 1997; Hildebrand, Dahlin, Volcani, 1998; Armbrust *et al.*, 2004; Thamatrakoln, Alverson, Hildebrand, 2006; Sapriel *et al.*, 2009). Проведенный нами филогенетический анализ показал, что пять белков SIT из *C. fusiformis* формируют отдельный кластер (**рис. 8**). Четыре из пяти SIT *P. tricornutum* также формируют отдельную группу, а один (PtSIT3) с высокой будструп-поддержкой кластеризуется с определенной в настоящей работе последовательностью SaSIT2. PtSIT3 и SaSIT2 оказались наиболее сходными и по профилям гидропатии аминокислотных последовательностей (**рис. 9**).

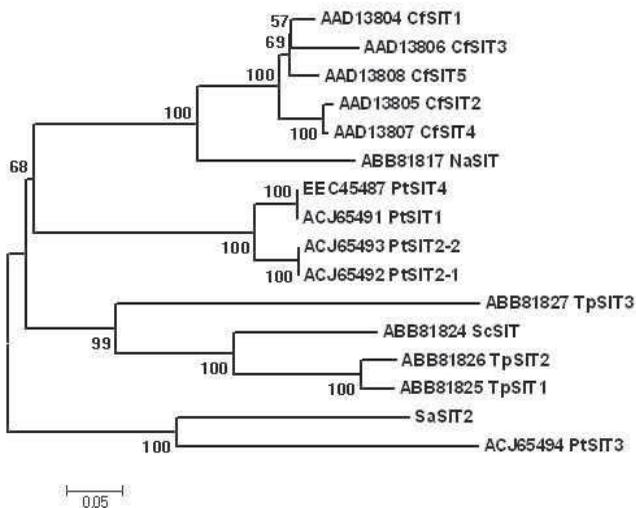


Рис. 8. Филогенетический анализ белков SIT методом ближайших соседей. В узлах дерева - бутстреп-значения, полученные при $n = 1000$ реплик. CfsIT – *C. fusiformis*; NaSIT – *N. alba*; PtsIT – *P. tricornutum*; ScSIT – *S. costatum*; SaSIT – *S. acus* subs. *radians*; TpsIT- *T. pseudonana*.

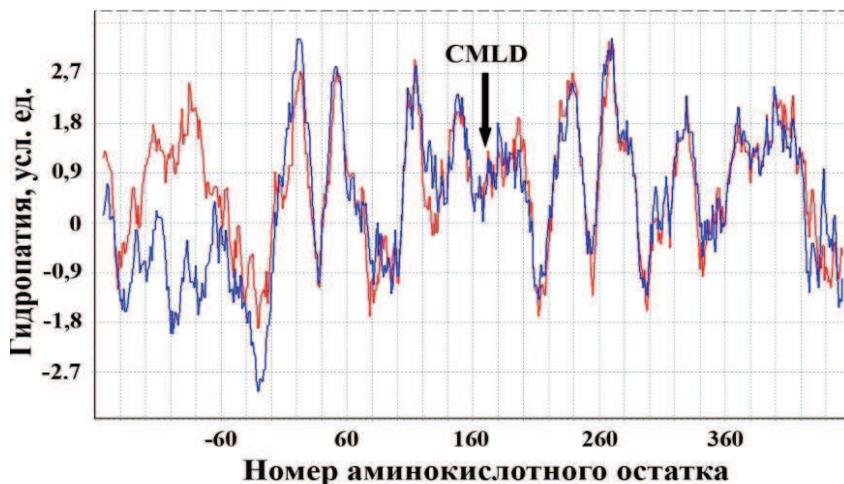


Рис. 9. Профили гидропатии предсказанных аминокислотных последовательностей SaSIT2 (красная линия) и PtsIT3 (синяя линия). Нумерация аминокислотных остатков дана по последовательности CfsIT1.

ВЫВОДЫ

1. Установлено, что профили гидропатии предсказанных аминокислотных последовательностей у всех 119 исследованных белков SIT из 58 видов диатомовых и из 2 видов хризофитовых водорослей сходны по расположению максимумов, минимумов и интенсивности пиков на перекрывающихся участках. Во всех белках SIT, за двумя исключениями, имеется консервативный мотив CMLD, либо его функциональный аналог.
2. Показано, что все белки SIT содержат протяженные участки длиной 42-110 а.о. без остатков лизина и аргинина. Эти, не содержащие положительно заряженных аминокислотных остатков, участки могут быть частью транспортной поры, по которой молекулы кремниевой кислоты свободно перемещаются из окружающей среды внутрь клетки.
3. Разработан метод получения суммарного белка из биомассы диатомей, основанный на растворении кремнистых панцирей в 5 М NH_4F при pH 5,0, который дает широкий спектр белков с молекулярными массами от 10 до 200 кДа.
4. На предсказанной аминокислотной последовательности белка SaSIT1 *Synedra acus* были выбраны четыре потенциальные антигенные детерминанты. Кроличьи антитела против синтетических пептидов, мимикрирующих эти детерминанты, позволили впервые для диатомей обнаружить белок SIT в протеоме методом иммуоблоттинга и выделить его из суммарного белка методом иммунохроматографии.
5. С использованием техники геномного пиросеквенирования у *S. acus* subsp. *radians* впервые обнаружен ген нового белка SIT - SaSIT2 - массой 62,4 кДа с 12 трансмембранными доменами, консервативным мотивом CMLD, входящим в состав протяженного участка (75 а.о.), не содержащего остатков лизина и аргинина. Профиль гидропатии белка SaSIT2 аналогичен профилям гидропатии других белков SIT.

Основные результаты диссертации опубликованы в следующих работах:

1. **Петрова Д. П.**, Калужная О. В., Костюковская А. О. Изучение растворения панцирей диатомовых водорослей *in vitro* // Бюлл. ВСНЦ РАН. –2004. – № 7. –С. 133-135.
2. Щербакова Т. А., Масюкова Ю. А., Сафонова Т. А., **Петрова Д. П.**, Верещагин А. Л., Минаева Т. В., Адельшин Р. В., Трибой Т. И., Стоник И. В., Козлов М. В., Лихошвай Е. В., Грачев М. А. Консервативный мотив CMLD в белках транспорта кремниевой кислоты у диатомовых водорослей // Молекулярная биология. –2005. –Т. 39. –С. 269-280.
3. Лихошвай Е. В., Масюкова Ю. А., Щербакова Т. А., **Петрова Д. П.**, Грачев М. А. Обнаружение гена транспорта кремниевой кислоты у хризодитовых водорослей // Доклады Академии наук. –2006. –Т. 408. –С. 1-5.
4. **Петрова Д. П.**, Бедошвили Е. Д., Шелухина И. В., Самуков В. В., Корнева Е. С., Верещагин А. Л., Попкова Т. П., Карпышев Н. Н., Лебедева Д. В., Клименков И. В., Лихошвай Е. В., Грачев М. А. Обнаружение белка транспорта кремниевой кислоты в клетках пресноводной диатомеи *Synedra acus* методами иммуноблоттинга и иммуноэлектронной микроскопии // Доклады Академии наук. –2007. –Т. 417. –С. 113-116.
5. Верещагин А. Л., Глызина О. Ю., Башарина Т. Н., Сафонова Т. А., Латышев С. А., Корнева Е. С., **Петрова Д. П.**, Анненков В. В., Даниловцева Е. Н., Чебыкин Е. П., Волокитина Н. А. Культивирование пресноводной диатомеи *Synedra acus* в столитровом фотобиореакторе и оценка состава биомассы // Биотехнология. –2008. –№ 4. –С. 55-63.
6. Shishlyannikov S. M., Zakharova Yu. R., Volokitina N. A., Mikhailov I. S., **Petrova D. P.**, Likhoshway Ye. V. A procedure for establishing an axenic culture of the diatom *Synedra acus* subsp. *radians* (Kütz.) Skabitsch. from Lake Baikal // Limnology and Oceanography: Methods. –2011. –V. 9. –P. 478-484.

Благодарности. Автор выражает благодарность акад. М.А. Грачеву за постановку задачи и внимание к работе, Ю.Р. Захаровой, С.М. Шишлянникову и Н.А. Волокитиной за предоставление биомассы клеток, Ю.П. Галачянцу за участие в пиросеквенировании, К.В. Хабудаеву за аннотирование генов, Ю.А. Масюковой за получение последовательностей генов *sit*, Ол.В. Калужной за совместную работу по отработке метода растворения панцирей диатомей, Г.И. Филипповой за подбор литературы, к.б.н. Т.А. Щербаковой и к.б.н. И.В. Клименкову за полезные советы и д.б.н. Е.В. Лихошвай за ценные замечания и помощь в оформлении работы. Работа выполнена при поддержке Программы Президиума РАН «Молекулярная и клеточная биология», проект 6.3.