

На правах рукописи

МЕДВЕДЕВ КИРИЛЛ ЕВГЕНЬЕВИЧ

**КОМПЬЮТЕРНОЕ ИССЛЕДОВАНИЕ ВЛИЯНИЯ ВЫСОКОГО
ДАВЛЕНИЯ И ТЕМПЕРАТУРЫ НА СТРУКТУРУ И ФУНКЦИЮ РНК-
СВЯЗЫВАЮЩЕГО БЕЛКА СЕМЕЙСТВА NP7 АРХЕЙ**

03.01.09 Математическая биология, биоинформатика

**Автореферат
диссертации на соискание ученой степени
кандидата биологических наук**

**Новосибирск
2014**

Работа выполнена в лаборатории эволюционной биоинформатики и теоретической генетики Федерального государственного бюджетного учреждения науки Институт цитологии и генетики СО РАН, г. Новосибирск, Россия.

Научный руководитель: кандидат биологических наук, доцент,
Афонников Дмитрий Аркадьевич

Официальные оппоненты: **Веселовский Александр Владимирович**
доктор биологических наук, заведующий лабораторией структурной биоинформатики, Федеральное государственное бюджетное научное учреждение Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича, г. Москва.

Бакулина Анастасия Юрьевна
кандидат биологических наук, старший научный сотрудник, теоретический отдел, Федеральное бюджетное учреждение науки Государственный научный центр вирусологии и биотехнологии «Вектор», Новосибирская область, р.п. Кольцово.

Ведущее учреждение: Федеральное бюджетное учреждение науки Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук, г. Москва.

Защита диссертации состоится «__» _____ 2015 г. на утреннем заседании диссертационного совета Д 003.011.01 по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук в ИЦиГ СО РАН в конференц-зале Института по адресу:
630090, г. Новосибирск, проспект ак. Лаврентьева, 10, т./ф. (383)363-49-06, факс (383) 333-12-78, e-mail: dissov@bionet.nsc.ru.

С диссертацией можно ознакомиться в библиотеке ИЦиГ СО РАН и на сайте Института www.bionet.nsc.ru.

Автореферат разослан «__» _____ 2014 г.

Ученый секретарь
диссертационного совета,
доктор биологических наук

Т.М. Хлебодарова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы

Высокие температуры и давления являются повреждающими факторами для живых клеток. Для человека и большинства известных нам организмов оптимальными являются условия температуры близкой к 27 С (300 К) и атмосферного давления (1 атм = 0.1 МПа). Однако на земле существуют организмы, которые заселяют местообитания с условиями, которые для человека являются экстремальными, например, температурой около 100 С (373 К) и давлениями выше 20 МПа (в 200 раз больше атмосферного). Такие условия наблюдаются вблизи глубоководных гидротермальных источников, заселенных сообществами микроорганизмов-экстремофилов (Rothschild & Mancinelli, 2001).

Механизмы, обеспечивающие выживание клеток в таких условиях остаются до конца не ясными. Их понимание позволит дать ответ на ряд фундаментальных вопросов, связанных с происхождением жизни и эволюцией микроорганизмов на ее ранних этапах (Daniel *et al.*, 2006) и адаптацией к условиям различных экологических ниш (Brooks *et al.*, 2011). В частности, актуальной задачей является выявление молекулярных механизмов эволюционной адаптации геномов и протеомов живых организмов к условиям высоких температур (Sternier and Liebl, 2001; Makarova *et al.*, 2003; Berezovsky *et al.*, 2007; McDonald, 2010; Jollivet *et al.*, 2012) и давлений (DiGuilio, 2005; Simonato *et al.*, 2006; Campanaro *et al.*, 2008). При этом одной из центральных фундаментальных задач в этих исследованиях является задача выявления механизмов адаптации структуры и функции белка к экстремальным условиям (прежде всего, к повышенным температурам и давлениям).

Немаловажен прикладной аспект этих исследований: их результаты могут быть использованы для создания ферментов, обладающих новыми свойствами, важными для биотехнологий (Podar & Reysenbach, 2006; Frock & Kelly, 2012). Ферменты такого рода востребованы в промышленности (Jaenicke *et al.*, 1998; Turner *et al.*, 2007; Scandurra *et al.*, 1998) так как химические реакции, протекающие при более высоких температурах, имеют более высокие скорости, что приводит к их большей эффективности. Кроме того, более термостабильные ферменты позволяют избежать нежелательных продуктов реакции, которые разрушаются при повышенных температурах (Vogt *et al.*, 1997). Термостабильные белки часто используются при биологическом анализе, например, в качестве биосенсоров. Для обеспечения повторяемости экспериментов с биосенсорами важно, чтобы ферменты как можно дольше сохраняли нативную структуру; также это позволяет избежать частой перекалибровки биосенсорных устройств. Кроме того, большое значение для фундаментальных исследований представляет оптимизация ДНК-полимераз (Talluri, 2011), которые используются в полимеразных цепных реакциях, в направлении термостабильности, специфичности и процессивности (Breyer *et al.*, 2001).

В последние годы, важную роль в понимании механизмов адаптации белков к экстремальным условиям среды наравне с экспериментальными подходами (см. обзор Liszka *et al.*, 2012) интенсивно используются методы биоинформатики и молекулярного моделирования. Биоинформатический анализ заключается, преимущественно, в сравнительном анализе белковых последовательностей и структур, относящихся к организмам с различными оптимальными условиями температур и давлений (Sternier and Liebl, 2001; Makarova *et al.*, 2003; Berezovsky *et*

al., 2007; Zeldovich *et al.*, 2007; McDonald, 2010; Jollivet *et al.*, 2012). Такое исследование интересно с биологической точки зрения и для понимания эволюционных аспектов реакций адаптивного ответа организмов на физиологический стресс. Вторым теоретическим направлением изучения механизма адаптации белков к экстремальным условиям, является использование методов молекулярного моделирования (Martinez *et al.*, 2011; Lee, 2011; Priyakumar *et al.*, 2010; Tiberti & Papaleo, 2011; Polyansky *et al.*, 2004; Calandrini & Kneller, 2008; Capese *et al.*, 2009; McCarthy & Grigera, 2006; Laurent *et al.*, 2012). Современные пакеты программ и вычислительные ресурсы позволяют проводить моделирование достаточно больших структур белков на временах десятков нс, что дает возможность детального исследования структурных характеристик моделей белков, как при нормальных, так и при экстремальных параметрах моделирования. Эти методы очень удобны при проведении сравнительного анализа динамических характеристик белков из организмов, обитающих в разных условиях. Они позволяют изучать на атомном уровне структурные различия мезофильных и экстремофильных белков, а также их изменение в процессе моделирования при разных давлениях и температурах. Эти преимущества методов молекулярного моделирования в последнее время привели к росту подобного рода сравнительных исследований, в том числе, влияния мутаций на стабильность белка.

Целью настоящей работы являлось изучение влияния высокого давления и температуры на пространственную структуру белка Nip7, участвующего в процессинге РНК у архей, методами молекулярной динамики и сравнительного анализа последовательностей и структур.

В соответствии с поставленной целью были сформулированы следующие задачи:

1. Методами молекулярной динамики провести моделирование структуры белка Nip7 мелководных и глубоководных архей рода *Pyrococcus* при атмосферном и повышенных давлениях, комнатной и высокой температурах;

2. Разработать подходы для анализа основных структурных характеристик компьютерных моделей белков Nip7 глубоководных и мелководных архей с целью выявления сходства и различий динамического поведения структур этих белков при разных параметрах температуры и давления на основе многофакторного дисперсионного анализа;

3. На основе анализа результатов моделирования выявить различия во флуктуациях доменов белка Nip7, охарактеризовать наиболее подвижные участки и выдвинуть предположение об их функции;

4. Провести поиск специфичных позиций в последовательностях семейства белков Nip7 архей, к таким параметрам среды местообитания как высокая температура и давление;

5. На основе результатов компьютерного анализа сформулировать гипотезу о возможных механизмах адаптации белков Nip7 к экстремальным условиям высоких давлений и температур.

Научная новизна и практическая значимость работы. Впервые проведено моделирование двух гомологичных белков из организмов, принадлежащих различным экологическим нишам при разных температурах (300, 373 К) и давлениях (0.1 – 100 МПа). Для сравнения структурных характеристик белков предложен метод двух и трехфакторного дисперсионного анализа.

Показано что структуры Nip7 *P.abyssi* и *P.furiosus* являются стабильными при всех параметрах температуры и давления, которые были исследованы, что свидетельствует о высокой устойчивости этих белков к экстремальным условиям температур и давлений.

На основе анализа компьютерных моделей белков Nip7 показано, что воздействие высокого давления и температуры на динамические свойства белков происходит неравномерно: наибольшие отклонения структур от кристаллографической модели и флуктуации полипептидной наблюдаются в районах петель; флуктуации С-терминального домена значимо выше, чем N-терминального.

На основе анализа аминокислотного состава, конформационной подвижности полипептидной цепи и компьютерного предсказания сайтов белок-белковых взаимодействий сформулирована гипотеза о вовлеченности аминокислот в позициях 49-59 (спирали α_2 - α_3) во взаимодействия Nip7 с белками экзосомы.

Выявлены участки полипептидной цепи моделей Nip7, конформационная динамика которых значимо отличается у белков мелководных и глубоководных организмов.

Выявлены позиции белка Nip7, в которых замены аминокислот ассоциированы с изменением температуры и глубины обитания организмов. Показано, что доля таких позиций, ассоциированных с температурой выше, чем с давлением.

Результаты работы позволяют лучше понять механизмы функционирования белков Nip7 архей и их возможной адаптации к экстремальным условиям. Полученные данные могут быть использованы при моделировании новых ферментов и их мишеней, функционирующих при высоких температурах и давлениях.

Положения, выносимые на защиту

1. Насыщенность структур белка Nip7 архей рода *Pyrococcus* солевыми мостиками является одним из механизмов его адаптации к экстремальным условиям повышенных температур и давления.

2. В исследованных интервалах значений высокие температуры оказывают более сильное влияние на изменение конформационных характеристик Nip7, чем высокие давления, а в процессе эволюции белков Nip7 их адаптация к высокотемпературным условиям среды более выражена, чем к глубоководным.

Апробация работы. Основные результаты работы были представлены на следующих конференциях: XLVIII Международная студенческая конференция «Студент и научно-технический прогресс», Новосибирск, 2010 г.; Первая Всероссийская молодёжная научная конференция, посвящённая 125-летию биологических исследований в Томском государственном университете «Фундаментальные и прикладные аспекты современной биологии», Томск, 2010г.; Российско-Индийский семинар "Предсказательная биология, интегративный анализ и методы", Индия, 2010г.; Международная Московская Конференция Вычислительной Молекулярной Биологии (МССМВ-11), Москва, 2011г.; Третья школа молодых ученых "Биоинформатика и системная биология", Новосибирск, 2011г.; II международная научно практическая конференция Постгеномные методы анализа в биологии, лабораторной и клинической медицине: геномика, протеомика, биоинформатика, 2011, Новосибирск; Восьмая Международная Конференция по Биоинформатике Регуляции Генома и

Структурной/Системной Биологии (BGRS/SB'2012), Новосибирск, 2012г.; IV Международная конференция «Математическая биология и биоинформатика» (ICMBB-2012), Пущино, 2012г.; Пятая Международная школа молодых ученых «Системная Биология и Биоинформатика» (SBB-2013), Новосибирск, 2013г.; Международная Московская Конференция Вычислительной Молекулярной Биологии (MCCMB-2013), Москва, 2013г.; Девятая Международная Конференция по Биоинформатике Регуляции Генома и Структурной/Системной Биологии (BGRS/SB-2014), Новосибирск, 2014г.

Публикации. По материалам диссертации опубликовано 3 работы, из них 2 работы – статьи в отечественной печати, в журналах, входящих в список ВАК.

Объем и структура работы. Диссертационная работа состоит из оглавления, списка сокращений, введения, обзора литературы, описания используемых материалов и методов, результатов, обсуждения, выводов, списка цитируемой литературы и 8 приложений. Работа изложена на 191 странице, содержит 22 рисунка и 6 таблиц. Библиографический указатель литературы включает 238 источников, из них 4 отечественных и 234 зарубежных.

Личный вклад автора. Основная часть работы выполнена автором самостоятельно. Основные результаты, представленные в публикациях, были получены автором.

Благодарности. Автор выражает глубокую признательность сотрудникам лаборатории эволюционной биоинформатики и теоретической генетики ИЦиГ СО РАН, и лично заведующему лабораторией и научному руководителю к.б.н. Афонникову Д.А., а также соавторам и коллегам – академику РАН Колчанову Н.А., к.б.н. Гунбину К.В., д.ф.-м.н. Воробьеву Ю.Н. (Лаборатория исследования модификации биополимеров, ИХБФМ СО РАН), д.х.н. Болдыревой Е.В. (Группа реакционной способности твердых веществ, ИХТТМ СО РАН) за консультации и плодотворные научные дискуссии. Автор благодарит д.ф.-м.н. Чекмарева С.Ф. (Лаборатория моделирования, ИТ СО РАН), к.б.н. Иванисенко В.А. (Лаборатория компьютерной протеомики, ИЦиГ СО РАН), к.ф.-м.н. Ломзова А.А. (Лаборатория бионанотехнологий, ИХБФМ СО РАН) за полезные критические замечания на стадии подготовки диссертации. Кроме того автор выражает благодарность к.б.н. Лашину С.А. (Лаборатория молекулярно-генетических систем, ИЦиГ СО РАН) и Алемасову Н.А. (Межинститутский сектор высокопроизводительных вычислений в биоинформатике, ИЦиГ СО РАН) за квалифицированную помощь и консультации при подготовке эксперимента по молекулярной динамике, а также при написании дополнительных программ на языке Java. Также автор выражает признательность Сибирскому Суперкомпьютерному Центру (ИВМиМГ СО РАН) и лично Кучину Н.В. за техническую поддержку в проведении экспериментов.

МАТЕРИАЛЫ И МЕТОДЫ

Модели белков. Для исследования изменений структурных свойств белка при повышенных давлениях были выбраны белок Nip7 *Pyrococcus abyssi* (пъезофильного организма, проживающего при давлениях около 20 МПа) и его гомолог у *Pyrococcus furiosus* (организма, при давлениях, близких к атмосферным). Эти белки были выбраны по следующим причинам. Во-первых, для них был идентифицирован режим движущего отбора в ходе эволюции архей *Pyrococcus* (Gunbin *et al.*, 2009). Следовательно, его функция может быть важной для

адаптации организмов к изменению давления. Во-вторых, для гомолога *Pyrococcus abyssi* известна пространственная структура (Coltri *et al.*, 2007, идентификатор PDB 2P38). В-третьих, длина его полипептидной цепи невелика, составляет 166 аминокислотных остатков, что приемлемо для моделирования структуры методом молекулярной динамики. В-четвертых, последовательности Nip7 *Pyrococcus abyssi* и его гомолога у *Pyrococcus furiosus* выравниваются без делеций/вставок, уровень сходства последовательностей составляет около 75%, что позволяет смоделировать пространственную структуру гомолога у *Pyrococcus furiosus* без реконструкции петель (этапа предсказания структуры белка по гомологии, который вносит существенные ошибки). Модель белка Nip7 *P.abyssii* была взята из банка данных PDB (Berman *et al.*, 2000), идентификационный номер 2P38 (Coltri *et al.*, 2007). Атомная модель белка Nip7 *P.abyssii* состоит из двух независимых мономеров и 252 молекул воды. Данная модель была обозначена как NIP7-ABY, нумерация аминокислотных остатков приводится согласно их порядковым номерам в структуре (1VAL – 155ASP).

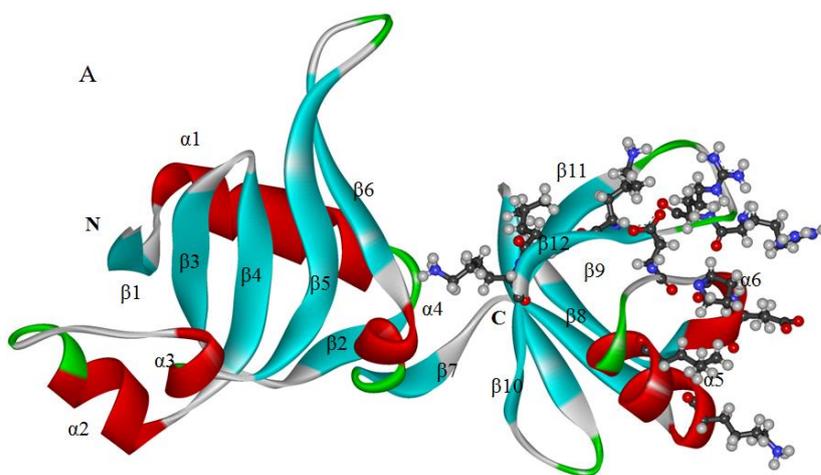


Рисунок 1.

Визуализация трехмерной структуры белка *P. abyssi* (2P38:A). Элементы вторичной структуры подписаны и окрашены соответствующим цветом (α -спирали красным, β -листы синим, петли зеленым). N-терминальный домен изображен слева, C-терминальный справа. Аминокислотные остатки, участвующие в связывании РНК (Coltri *et al.*, 2007)

изображены в виде шариков.

Мономеры белка Nip7 составлены из двух альфа-бета доменов. Массивный N-терминальный домен состоит из пяти антипараллельных β -листов, окруженных тремя α -спиралями и одной спиралью 3_{10} . Его функцией является взаимодействие с экзосомой. С-терминальный домен, который является РНК-связывающим доменом, включает аминокислотные остатки с 95 по 159 и состоит из смешанных β -листов, одной α -спирали и одной короткой спирали 3_{10} (Coltri *et al.*, 2007).

Модель NIP7 *P.furiosus* (NIP7-FUR), полученная методом реконструкции по гомологии, была взята из банка данных ModBase (Pieper *et al.*, 2011), ID Q8TZP7. Число остатков в модели NIP7-FUR равно 155, их нумерация соответствует модели NIP7-ABY (1ILE-155ASP). Среднеквадратичное отклонение моделей NIP7-FUR и NIP7-ABY по координатам C^α составило 0.1 Å.

В работе исследовалась молекулярная динамика белков при давлениях 0,1, 50, 100 МПа и температурах 300 и 373 К. Для каждого набора параметров давления и температуры осуществлялось пять запусков молекулярной динамики. **Моделирование** осуществлялось с помощью пакета программ GROMACS 4.5.3 (Van Der Spoel *et al.*, 2005) с использованием алгоритма LINCS (Hess *et al.*, 1997) в силовом поле amber99sb-ildn (Berhanu *et al.*, 2012); электростатические взаимодействия рассчитывались с помощью алгоритма Partial Mesh Ewald (PME).

Время моделирования равнялось 40 наносекундам. Данные сохранялись каждые 100 пс.

Сравнение моделей и оценка подвижности полипептидной цепи

Мы анализировали такие структурные характеристики белков как радиус гирации (R_g), площадь поверхности остатков доступной растворителю (SAS), среднеквадратичное отклонение координат C_α атомов от кристаллографической структуры Nip7 *P.furiosus* (RMSD), среднеквадратичные флуктуации C_α атомов (RMSF), вторичную структуру (Kabsch *et al.*, 1983), солевые мостики (Barlow *et al.*, 1983) а также локальное отклонение структуры – RMSDL в скользящем окне длиной $q+1$ для каждого остатка:

$$RMSDL_i(\mathbf{v}, \mathbf{w}, q) = \sqrt{\left(\frac{1}{q} \sum_{j=i-q/2}^{j=i+q/2} (v_{jx} - w_{jx})^2 + (v_{jy} - w_{jy})^2 + (v_{jz} - w_{jz})^2 \right)},$$

где \mathbf{v} , \mathbf{w} – аминокислотные последовательности, v , w – конкретные аминокислотные остатки этих последовательностей, x , y , z – координаты каждого аминокислотного остатка, n – общее количество остатков, q – движущееся окно размером 11 аминокислотных остатков.

Для того чтобы определить степень влияния повышенной температуры (давления) на конформационные параметры белка, был введён индекс чувствительности S^t . Были рассчитаны средние значения конформационного параметра X для траекторий при высокой ($X_{T=373}$) и низкой ($X_{T=300}$) температуре. Индекс чувствительности рассчитывался следующим образом: $S^t(S^P) = \frac{X_{T(P)=373}}{X_{T(P)=300}}$.

Если значение $S^t(S^P)$ близко к единице это означает, что высокая температура (давление) не имеет значительного влияния на данный конформационный параметр по сравнению с его значением при низкой температуре (давлении).

Для того чтобы оценить влияние температуры, давления и аминокислотных замен на **конформационные параметры** белков мы использовали два типа дисперсионного анализа ANOVA (Kendall *et al.*, 1968).

Для того чтобы учесть влияние температуры и давления на общие характеристики белковой глобулы, мы использовали двухфакторный дисперсионный анализ. При анализе структурных характеристик аминокислотных остатков (RMSDL и RMSF) мы дополнительно оценивали влияния типа белка (NIP7-ABY, NIP7-FUR), рассматривая, таким образом, модель трехфакторного дисперсионного анализа с учетом температуры, давления и типа белка.

Для **визуализации моделей структур** использовалась программа Accelrys Discovery Studio Visualizer. **Статистическая обработка** данных была проведена с помощью программ Excel и Statistica 6.0. Для **предсказания сайтов белок-белковых** взаимодействий использовался сервер SPPIDER (Porollo *et al.*, 2007)

Для **анализа эволюции белков Nip7 архей** была составлена выборка, включающая последовательности его гомологов 35 организмов. Последовательности были выровнены с использованием программы PROMALS3D (Pei *et al.*, 2008). Мы использовали для идентификации специфичных замен две программы, multi-Harmony (Pirovano *et al.*, 2006) и GroupSim (Capra *et al.*, 2008).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Молекулярное моделирование белковых структур Nip7 глубоководного и мелководного организмов

Стабильность моделей белков

Из Таблицы 1 видно, что отклонения полипептидной цепи моделей от кристаллографической структуры увеличиваются с повышением температуры и давления. Однако, среднее значение RMSD не превышает 2.4 Å, что свидетельствует о стабильности обеих моделей моделирования при всех параметрах внешней среды.

Таблица 1. Средние значения RMSD двух моделей белков NIP7-ABY и NIP7-FUR относительно кристаллографической структуры Nip7 *P. abyssi* (2P38:A) и их 95%-доверительные интервалы при различных значениях давления и температуры.

Модель	NIP7-ABY		NIP7-FUR	
	Температура 300 К	Температура 373 К	Температура 300 К	Температура 373 К
0.1	1.76 ± 0.04	1.93 ± 0.04	2.08 ± 0.03	2.41 ± 0.02
50	1.81 ± 0.04	1.94 ± 0.06	2.2 ± 0.05	2.2 ± 0.03
100	1.79 ± 0.06	1.78 ± 0.03	2.18 ± 0.02	2.23 ± 0.02

У модели NIP7-FUR наблюдаются бóльшие значения RMSD, чем у NIP7-ABY при всех значениях давления и температуры. Таблица демонстрирует также, что при увеличении температуры наблюдается тенденция увеличения RMSD для двух белков.

На Рисунке 2 показано сравнение средних значений радиусов гирации (Rg) моделей NIP7-ABY и NIP7-FUR при разных давлениях и температурах.

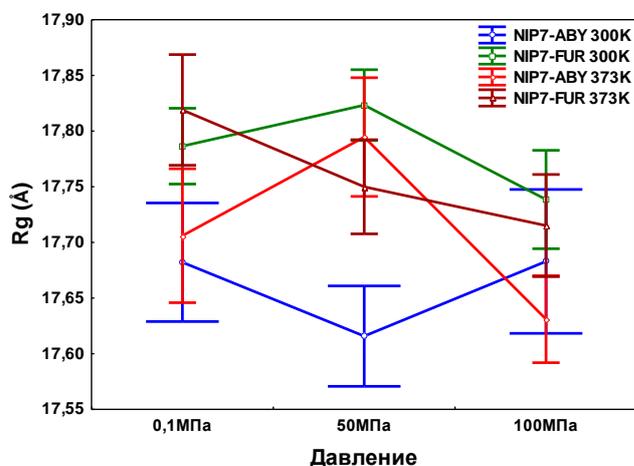


Рисунок 2. Зависимость значения радиуса гирации (Rg, ось Y) от давления (ось X) при высокой и низкой температурах для двух моделей белка Nip7 NIP7-ABY и NIP7-FUR. Вертикальные линии соответствуют 95% доверительным интервалам.

Рис. 2 демонстрирует, что значение Rg глубоководного организма NIP7-ABY меньше, то есть этот белок имеет более плотную упаковку, чем её мелководных гомолог при давлении 0,1 МПа и 100

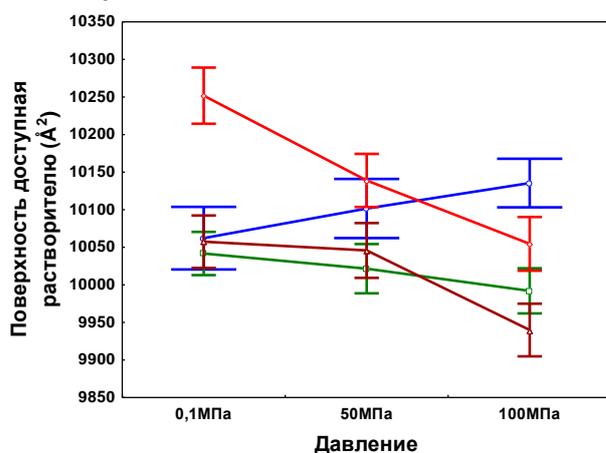
МПа при любом значении температуры. Двухфакторный анализ величин Rg для двух моделей показал, что изменения давления и температуры не приводят к значимым изменениям средних значений их Rg (привести тут значение статистик).

Взаимодействие с растворителем

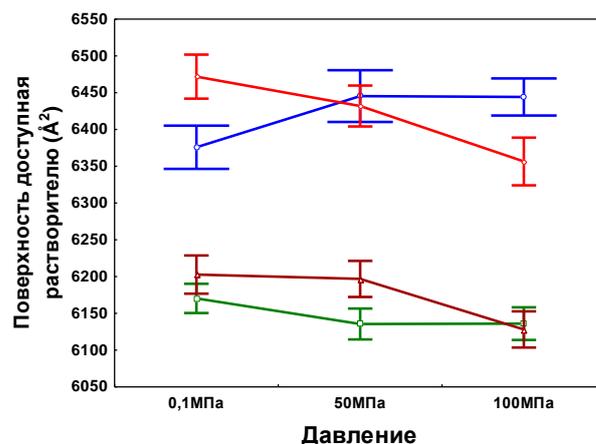
Для каждой траектории двух моделей белков мы рассчитали доступность остатков растворителю для всего белка (SAS_t), также отдельно полярную (SAS_p) и гидрофобную (SAS_h) части поверхности. На Рисунке 3 (А-В) мы сравнили параметры SAS_t , SAS_p , и SAS_h полученные для различных траекторий. Как видно из Рисунка 3А, относительная разница в полной площади поверхности двух белков находится в пределах стандартной ошибки. Этот факт согласуется с разницей в

значениях радиуса гирации, приведённой выше. Общая площадь поверхности несколько меньше у NIP7-FUR.

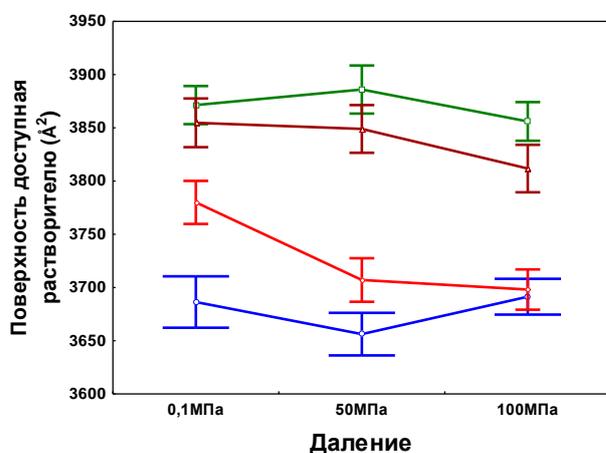
А SAS_t



Б SAS_p



В SAS_h



■ NIP7-ABY 300K

 ■ NIP7-FUR 300K

 ■ NIP7-ABY 373K

 ■ NIP7-FUR 373K

Рисунок 3. Графики зависимости площади поверхности белка доступной растворителю от различных условий моделирования молекулярной динамики: (А) общая поверхность белка доступная растворителю; (Б) полярная часть; (В) гидрофобная часть. Давление отложено вдоль оси X, параметр SAS вдоль оси Y. Вертикальные линии отражают 95% доверительные интервалы.

Из рисунка 3 видно, что для модели NIP7-ABY наблюдаются более высокие значения полярной доли остатков, доступных растворителю, чем для модели NIP7-FUR. Площадь гидрофобной части поверхности для этой модели меньше. Эти различия наблюдаются для всех исследованных нами параметров температуры и давления.

Мы оценили влияние температуры и давления на изменение площади поверхностей моделей на основе двухфакторного дисперсионного анализа (Таблица 2). Полученные данные в целом согласуются с анализом по Рисунку 3.

Температура значимо влияет на изменение гидрофобной части площади поверхности модели NIP7-ABY, при этом показано, что влияние давления и температуры на изменение полной площади поверхности белка не является аддитивным. Это, по-видимому, отражает различие в значениях SAS для высоко- и низкотемпературных траекторий NIP7-ABY. Для модели NIP7-FUR показано, что давление значимо влияет на изменения как гидрофильной части площади, так и SAS_t. Таким образом, давление и температура по разному влияют на изменение

доступности остатков растворителю для двух моделей белков. Давление приводит к уменьшению площади поверхности, доступной растворителю. Это наиболее выражено для моделей при высоких температурах.

Таблица 2. Значение F-статистики для поверхности белка доступной растворителю. SAS_t – полная поверхность белка, SAS_h – гидрофобная, SAS_p – полярная. Величины, превышающие критические значения при 5% уровне значимости выделены жирным шрифтом.

Фактор внешней среды	NIP7-ABY			NIP7-FUR		
	SAS_t	SAS_h	SAS_p	SAS_t	SAS_h	SAS_p
Температура	0,93	2,23	0,51	5,03	1,67	4,04
Давление	1,72	5,94	0,003	0,03	4,01	3,27
Взаимодействие	4,46	1,47	2,92	1,12	0,26	1,61

Локальная структура полипептидной цепи

Анализ чувствительности значения RMSDL каждого аминокислотного остатка к увеличению давления (S_i^p), температуры (S_i^t) показал, что локальные структурные отклонения полипептидной цепи от кристаллографической структуры сильно отличаются для разных аминокислотных остатков (Рис. 4).

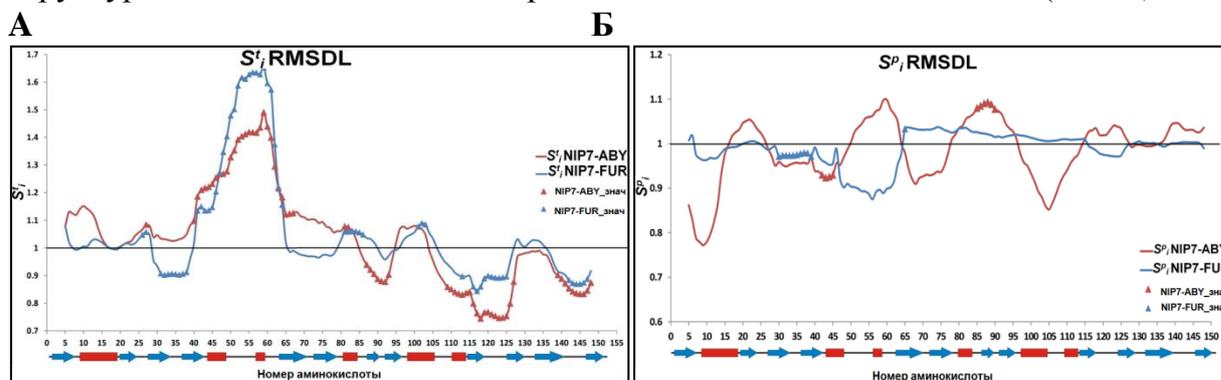


Рисунок 4. Графики зависимости индекса чувствительности RMSDL к увеличению температуры (А) и давления (Б) от номера аминокислотного остатка. Треугольниками показаны значимые отклонения по результатам двухфакторного дисперсионного анализа.

Во-первых, существуют участки белка, для которых температура значимо увеличивает RMSDL ($S_i^t > 1$) для двух моделей (Рис. 4А). К таким районам относится, прежде всего, большой сегмент последовательности от С-терминальной части нити β_4 до N-терминального конца нити β_5 (позиции 41-65). Для С-терминального домена подавляющая часть остатков имеет значения $S_i^t < 1$, для многих из них эти отличия значимы. Это означает, что повышение температуры делает конформацию С-терминального домена в большей степени похожей на конформацию кристаллографической структуры. Наиболее существенны эти изменения для участка 109-127 ($\alpha_6 - \beta_{10}$). Значимое уменьшение RMSDL характерно и для С- терминальных остатков в обеих моделях. Чувствительность параметра RMSDL к увеличению давления показана на Рисунке 4Б. Для модели NIP7-ABY отклонения значений S_i^p вдоль последовательности более выражены, однако значимыми являются изменения в позициях 41-44 (уменьшение RMSDL) и 85-89 (увеличение RMSDL).

Флуктуации полипептидной цепи

Индексы чувствительности RMSF остатков к повышению температуры и давления для двух моделей показаны на Рисунке 5. Из рисунка видно, что

повышение температуры увеличивает значение локальных полипептидных флуктуаций (S_i^t больше единицы для подавляющего числа остатков; рис 5А). Остатки, чувствительность которых к повышению температуры мала (NIP7-ABY: 36, 121, 123, 131; NIP7-FUR: 146) располагаются в петлях и имеют высокие значения RMSF уже при температуре 300К, т.о. увеличение температуры не приводит к какому-либо существенному увеличению их подвижности.

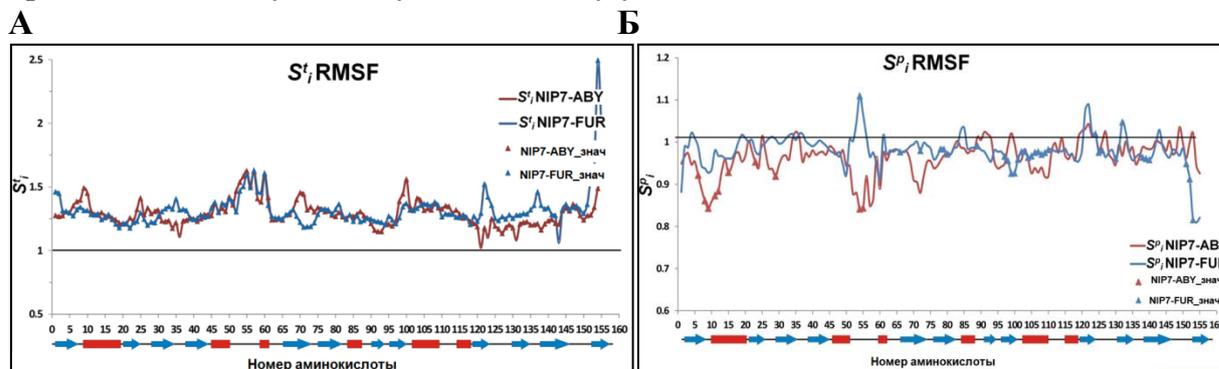


Рисунок 5. Графики зависимости индекса чувствительности RMSF к увеличению температуры (А) и давления (Б) от номера аминокислотного остатка. Треугольниками показаны значимые отклонения по результатам двухфакторного дисперсионного анализа

Индексы чувствительности RMSF к давлению отличаются для двух моделей в большей степени, чем для температур (Рисунок 5Б). S_i^p превышает единицу только для небольшого числа остатков. Для N-терминального домена уменьшение флуктуаций для модели NIP7-ABY при воздействии давления выражено больше, чем для NIP7-FUR. В этом домене для ряда остатков давление значимо влияет на изменение RMSF.

Влияние температуры и давления на флуктуации остатков N- и С-терминального доменов двух моделей сравнили при помощи двухфакторного анализа (табл. 3)

Таблица 3. Значение F-статистики для параметра RMSF двух доменов белка Nip7. Величины, превышающие критические значения при 5% уровне значимости выделены жирным шрифтом

Факторы	NIP7-ABY		NIP7-FUR	
	300К	373К	300К	373К
Давление	0,44	4,27	4,92	3,68
Домены	7,88	10,04	5,39	28,92
Взаимодействие	0,42	0,54	0,06	5,05

Статистический анализ показал, что для модели NIP7-ABY при температуре 300К изменение давления не влияет значимо на флуктуации N- и С-терминального домена. Однако, различия во флуктуациях между доменами доменов значимы: С-терминальный домен демонстрирует большие флуктуации по сравнению с N-терминальным доменом. При высокой температуре влияние давления на параметр RMSF двух доменов становится значимым. Для модели NIP7-FUR наблюдается значимое влияние давления на величину параметра RMSF обоих доменов при низком давлении и высокой температуре. Флуктуации обоих доменов снижаются с увеличением давления, как и в модели NIP7-ABY. Различия в значениях RMSF были значимы при температуре 300 и 373К.

Анализ солевых мостиков

Солевые мостики в структуре моделей NIP7-ABY и NIP7-FUR были разделены по их персистентности (доли структур в которых они наблюдаются) на стабильные (в среднем более 70% структур), умеренно стабильные (от 70 до 20% структур) и неустойчивые (менее 20%). Шесть солевых мостиков (GLU10-ARG4, GLU75-ARG37, ARG37-GLU10, LYS20-GLU17, ARG148-ASP109, GLU33-ARG4) являются наиболее стабильными (присутствуют более чем в 70% структур, Таблица 4). Четыре стабильных солевых мостика, расположенных в N-терминальном домене (GLU10-ARG4, GLU75-ARG37, ARG37-GLU10, GLU33-ARG4) формируют сеть, соединяющую альфа-спираль с бета-листами. Большая часть солевых мостиков встречается в обоих моделях: все стабильные и умеренно стабильные мостики у NIP7-FUR являются таковыми и у NIP7-ABY. Их число больше у модели NIP7-ABY (18), чем у NIP7-FUR (11).

Таблица 4. Чувствительность наиболее стабильных солевых мостиков в моделях NIP7-ABY и NIP7-FUR повышению температуры и давления. Первая колонка – номер аминокислот, формирующих солевой мостик; вторая колонка – наличие солевых мостиков в разных структурах (ALL: модели NIP7-ABY, NIP7-FUR и кристаллографическая структура 2P38:A; ABY & FUR: только для моделей NIP7-ABY, NIP7-FUR и т.д.); третья колонка – указание расположение солевого мостика в структуре белка; колонки 4-11 – значения индексов чувствительности для моделей NIP7-ABY и NIP7-FUR. F – значение F-критерия для конкретного параметра по результатам двухфакторного дисперсионного анализа. Значимые значения выделены жирным шрифтом. Нестабильные солевые мостики (процент встречаемости менее 20%) отмечены символом – н.

Мостик	Наличие	Расположение	<i>P.abysssi</i>				<i>P.furiosus</i>			
			S^T	$F(T)$	S^P	$F(P)$	S^T	$F(T)$	S^P	$F(P)$
10-4	все	$\alpha_1 - \alpha_1$	0,94	2,74	1,04	0,83	0,99	1	1,01	1
75-37	все	$\beta_6 - \beta_4$	1,06	0,52	1,11	0,91	0,99	1,89	1,00	0,36
37-10	все	$\beta_4 - \alpha_1$	1,04	0,16	1,14	0,99	1,18	2,87	0,88	4,68
148-109	все	РНК связывающий - РНК связывающий	0,95	0,40	1,09	0,45	н	н	н	н
33-4	все	$\beta_3 - \beta_1$	0,92	0,59	0,94	0,69	0,98	0,10	1,00	0,20
20-17	все	$\alpha_1 - \alpha_1$	0,92	2,65	1,02	0,13	0,89	2,96	0,98	0,36
16-12	P.a.&P.f.	$\alpha_1 - \alpha_1$	1,14	3,77	0,91	1,68	1,08	1,60	0,89	2,76
147-113	P.a.&P.f.	РНК связывающий - РНК связывающий; α_6	0,98	0,02	0,72	3,18	н	н	н	н
33-2	2p38&P.a.	β_3 - N-конец	0,98	0,01	0,99	2,07	-	-	-	-
131-116	все	C-конец - β_9	1,93	12,23	0,94	0,49	0,98	0,72	0,99	0,43
51-5	P.a.&P.f.	$\alpha_2 - \beta_1$	1,25	1,50	1,03	0,03	0,88	1,11	1,18	0,74
76-70	P.a.&P.f.	β_6 - N-конец	1,20	6,59	0,84	3,42	0,83	8,19	0,95	1,98
147-146	P.a.	РНК связывающий - C-конец	1,84	4,12	0,95	0,12	-	-	-	-
48-45	P.a.&P.f.	$\alpha_2 - \alpha_2$	1,06	0,14	0,92	1,42	0,99	0,00	0,88	1,29
88-23	P.a.	$\beta_7 - \beta_2$	1,26	5,91	0,87	1,72	-	-	-	-
151-109	все	РНК связывающий; β_{12} - РНК связывающий	0,93	0,09	0,75	1,49	0,86	2,74	0,92	0,42
119-117	P.a.	$\beta_9 - \beta_9$	0,98	0,04	1,26	2,31	-	-	-	-
48-44	2p38&P.a.	$\alpha_2 - \alpha_2$	0,99	0,00	0,98	0,01	-	-	-	-
146-112	все	C-конец - α_5	н	н	н	н	1,01	0,00	1,11	0,50
76-17	P.a.&P.f.	$\beta_6 - \alpha_1$	н	н	н	н	0,83	1,07	0,95	0,56

Анализ функциональной роли наиболее стабильных солевых мостиков показал (Таблица 4), что большинство из них стабилизируют упаковку элементов вторичной структуры белка.

Результаты дисперсионного анализа персистентности для стабильных и умеренно стабильных солевых мостиков в обеих моделях представлены в Таблице 4. Они демонстрируют, что для модели NIP7-ABY 3 из 20 стабильных солевых (17%) повержены влиянию высокой температуры (131-116, 76-70 и 88-23). Два солевых мостика значимо подвержены влиянию температуры (148-109 и 76-70), оба они также стабилизируются при высокой температуре. При этом первый из них сформирован остатками, участвующими в связывании РНК, а второй также стабилизируется при высоких температурах у NIP7-ABY. Таким образом, повышение температуры для ряда солевых мостиков в наших моделях изменяет их персистентность, преимущественно в сторону увеличения.

Влияние типа белка на динамику белковой структуры

Результаты оценки значимости влияния типа модели (NIP7-ABY или NIP7-FUR) на параметры RMSDL и RMSF остатков показаны на рис. 6. Тип белка значимо влияет на локальное отклонение полипептидной цепи от кристаллографической структуры (2P38:A) для 77 из 143 проанализированных нами позиций (53%). Давление оказывает значимое влияние на RMSDL только для 9 позиций белка (6%). В тоже время тип белка значимо влияет на RMSDL остатков для 91 позиции (63%), Рисунок 6.

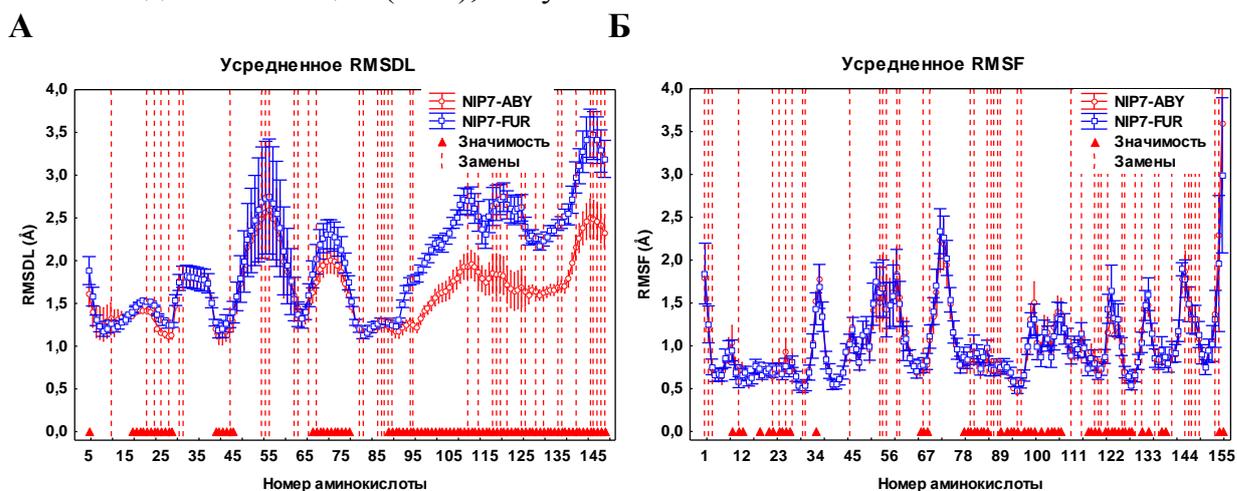


Рисунок 6. Усредненные значения RMSDL (А) и RMSF (Б) для белков NIP7-ABY и NIP7-FUR по всем траекториям и запускам. Красные треугольники показывают позиции, в которых параметр RMSDL/RMSF значимо зависит от типа модели. Красные пунктирные линии – позиции замен в NIP7-ABY относительно NIP7-FUR.

Наиболее характерные различия наблюдаются для С- терминального домена белка, а в N-терминальном домене и в позициях 5,16-28, 40-45, 67-77. Для этих участков отклонения от конформации кристаллографической структуры у модели NIP7-FUR выше, чем у модели NIP7-ABY.

Результаты аналогичного анализа для RMSF аминокислотных остатков показали, что повышение температуры значимо влияет на увеличение RMSF для всех без исключения аминокислотных остатков белка NIP7. Давление значимо влияет на изменение RMSF у 26 аминокислотных остатков (16%). Тип модели влияет на изменение флуктуаций для 56 (36%) аминокислотных остатков

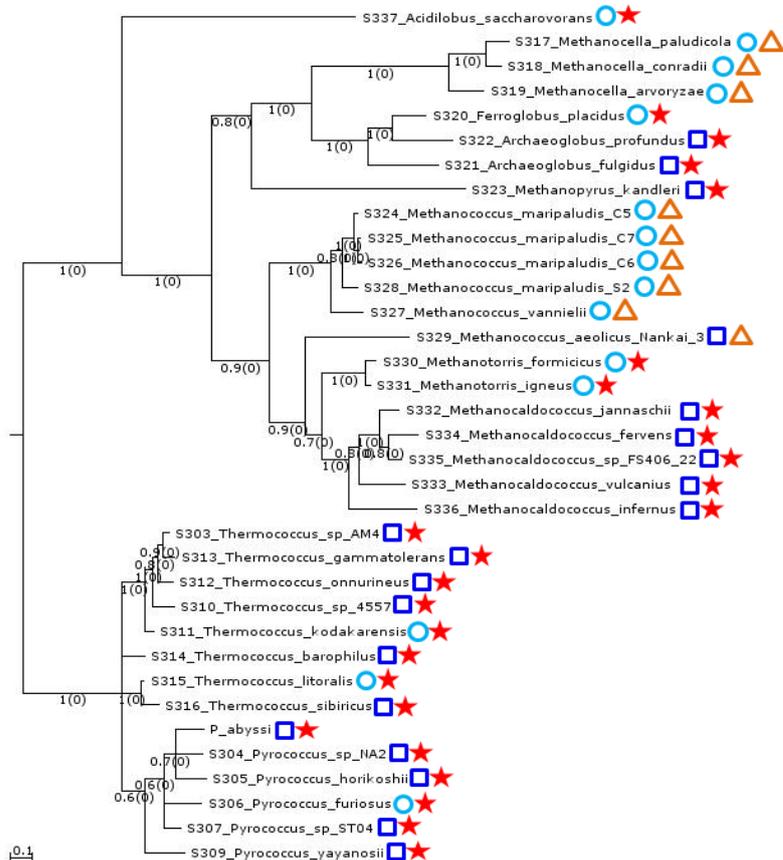
(Рисунок 6Б). Значимые различия между значениями RMSF для остатков из двух моделей наблюдались как в случае, когда флуктуации были выше для модели NIP7-ABY, так и, наоборот, для NIP7-FUR (Рисунок 6Б).

Поиск мутаций, связанных с адаптацией к экстремальным условиям в белках Nip7 архей. Анализ координированных замен

Для выявления позиций, замены в которых происходят специфическим образом по отношению к условиям внешней среды организмов, была собрана выборка из 35 белков гомологов Nip7 архей. В данную выборку вошли белки представителей следующих родов архей: *Pyrococcus*, *Thermococcus*, *Methanococcus*, *Methanocaldococcus*, *Methanocella*, *Ferroglobus*, *Archaeoglobus*, *Methanopyrus*, *Methanotorris*, *Acidilobus*. Полученная выборка была разделена по каждому из двух экологических признаков организмов (глубина и температура местообитаний) на две подгруппы: а) мелководные (глубина менее 1000 м) и глубоководные организмы (глубина более 1000 м), б) термофильные (температура ниже 45°C) и термофильные организмы (температура выше 45°C).

На рисунке 6 представлено филогенетическое дерево, построенное по выравниванию белков Nip7 архей из данной выборки. На нем для каждого вида показано, к какой экологической группе он относится. Полученное дерево в целом адекватно отражает таксономическое деление изученных организмов. Все они, за исключением *M. aeolicus*, группируются согласно родам. Анализ филогении показывает, что представители одного и того же рода обитают, преимущественно, в схожих условиях. В отношении к термофильности эта тенденция выражена в большей степени, чем в отношении глубин местообитаний. Так, без исключений, термофильными являются роды *Thermococcus*, *Pyrococcus* и *Methanocaldococcus*.

Результаты анализа множественного выравнивания программами GroupSim



(Capra *et al.*, 2008) и multi-Harmony (Pirovano *et al.*, 2006) приведены на Рисунке 7. На данном рисунке приведены трехмерные структуры белка Nip7, на которых отмечены позиции специфичные: А – к давлению, Б – к температуре, по совокупному результату работы двух программ.

Рисунок 6. Филогенетическое дерево последовательностей белков Nip7 выборки архей. Условия обитания определены следующими условными обозначениями: синий квадрат – глубоководный организм, голубой круг – мелководный организм, красная звезда – термофильный организм, оранжевый треугольник – мезофильный организм.

4. На основе анализа гомологичных последовательностей белка Nip7 архей, выявлены позиции, замены в которых происходят специфическим образом по отношению к температуре и давлениям в местообитаниях этих организмов.

5. Согласно результатам проведенного анализа, одним из факторов стабилизации структуры Nip7 являются сети солевых мостиков, количество которых больше у белков глубоководных и термофильных организмов; аминокислотные остатки на поверхности N-терминального домена у глубоководных организмов более полярны, чем у мелководных, что может быть связано с особенностями их возможных взаимодействий с субъединицами экзосомы архей, которая отвечает за деградацию молекул РНК.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

1. **Медведев К.Е.**, Афонников Д.А., Воробьев Ю.Н. Изучение методами молекулярной динамики структур белков NIP7 глубоководных и мелководных архей в условиях повышенного давления // Вестник Томского государственного университета. Биология № 4 (16), с136, 2011
2. Афонников Д.А., **Медведев К.Е.**, Гунбин К.В., Колчанов Н.А. Важная роль гидрофобных взаимодействий при адаптации белков к высоким давлениям // Доклады Академии Наук 438(3):412–415, 2011
3. **Медведев К.Е.**, Афонников Д.А. Анализ структур белков архей-пъезофилов под влиянием высокого давления с помощью компьютерных методов // Труды Томского государственного университета, сер. Биологическая. Т. 275, С. 378-380, 2011.
4. **К.Е. Медведев.** Анализ структур белков в условиях повышенных давлений при помощи компьютерных методов // XLVIII Международная студенческая конференция «Студент и научно-технический прогресс», Новосибирск, 10 – 14 апреля 2010 г., с.4.
5. **К.Е. Медведев**, Д.А. Афонников. Анализ структур белков архей-пъезофилов под влиянием высокого давления с помощью компьютерных методов // Первая Всероссийская молодёжная научная конференция, посвящённая 125-летию биологических исследований в Томском государственном университете «Фундаментальные и прикладные аспекты современной биологии», Томск, 6 – 9 октября 2010г., с.132.
6. Dmitry A. Afonnikov, Konstantin V. Gunbin, **Kirill E. Medvedev** Prokaryotic adaptation to different environmental conditions at the genomic level: Molecular evolution analysis // Joint Indo-Russian Workshop «Predictive Biology using Systems and Integrative analysis and Methods» November 15-19.2010, Chandigarh, India
7. **К.Е. Medvedev**, D.A.Afonnikov Molecular dynamics simulation of Nip7 proteins demonstrates importance of hydrophobic interaction for protein stability at high pressure // International Moscow Conference on Computational Molecular Biology (MCCMB 11)
8. **Медведев К.Е.** Исследование влияния высокого давления и высокой температуры на структуру белка Nip7 мелководных и глубоководных архей рода *Rugosoccus* методами молекулярной динамики // Третья школа молодых ученых "Биоинформатика и системная биология" 2011г.
9. **Медведев К.Е.**, Афонников Д.А. Исследование структурных изменений белка Nip7 мелководных и глубоководных архей рода *Rugosoccus* под влиянием высокого давления и высокой температуры методами молекулярной динамики //

II международная научно практическая конференция Постгеномные методы анализа в биологии, лабораторной и клинической медицине: геномика, протеомика, биоинформатика 2011г.

10. **К.Е. Medvedev**, D.A. Afonnikov, Y.N. Vorobjev. Molecular dynamics simulation of nip7 proteins from hyperthermophilic archaea at high temperature and pressure // BGRS/SB'2012.
11. **Медведев К.Е.**, Афонников Д.А., Воробьев Ю.Н. Исследование влияния высокого давления и высокой температуры на структуру белка Nip7 архей рода *Rugosoccus* методами молекулярной динамики // IV Международная конференция «Математическая биология и биоинформатика» (ICMBB-2012)
12. **Medvedev К.Е.**, Afonnikov D.A., Vorobjev Y.N. Molecular dynamics simulation analysis of influence of high temperature and high pressure on the nip7 proteins from the hyperthermophilic archaea // Systems Biology & Bioinformatics 2013
13. **К.Е. Medvedev**, K.V. Gunbin, Y.N. Vorobjev, D.A. Afonnikov. Influence of high temperature and high pressure on the Nip7 proteins from the hyperthermophilic archaea *P. abyssi* and *P. furiosus*: molecular dynamics simulation analysis // MCCMB-2013