

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ
СОЕДИНЕНИЙ

УДК 539.26

ИЗМЕНЕНИЯ ЧЕТВЕРТИЧНОЙ СТРУКТУРЫ
НУКЛЕОТИД-РЕГУЛИРУЕМОЙ ПИРОФОСФАТАЗЫ
ИЗ *Desulfitobacterium hafniense* ПРИ ВЗАИМОДЕЙСТВИИ С ЛИГАНДАМИ,
ВЫЯВЛЕННЫЕ МЕТОДОМ МАЛОУГЛОВОГО РЕНТГЕНОВСКОГО
РАССЕЯНИЯ В РАСТВОРЕ

© 2020 г. Л. А. Дадинова^{1,*}, В. А. Анашкин², Э. В. Штыкова¹

¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Научно-исследовательский институт физико-химической биологии им. А.Н. Белозерского Московского государственного университета им. М.В. Ломоносова, Москва, Россия

*E-mail: lbovmsu@mail.ru

Поступила в редакцию 28.02.2020 г.

После доработки 28.02.2020 г.

Принята к публикации 03.03.2020 г.

Для установления механизма передачи сигнала между активным и регуляторным центрами регулируемой нуклеотидами неорганической пирофосфатазы интенсивно исследуется регуляция этих ферментов на молекулярном уровне. Однако отсутствие достоверных структурных данных о полномасштабном белке и его взаимодействии с лигандами не позволяет однозначно решить проблему. Методом малоуглового рентгеновского рассеяния впервые определена структура низкого разрешения регулируемой нуклеотидами пирофосфатазы из *Desulfitobacterium hafniense*, а также выявлены пространственные изменения полномасштабного фермента при его взаимодействии с аденозинмонофосфатом и диаденозинтетрафосфатом. Показано, что белок в разбавленных растворах существует как стабильный гомотетрамер, структура которого зависит от природы присоединенного лиганда. Полученные структурные результаты важны для понимания молекулярных основ регуляции ферментов этого семейства.

DOI: 10.31857/S0023476120050057

ВВЕДЕНИЕ

Неорганические пирофосфатазы (РРазы) — ферменты, присутствующие во всех живых организмах. Они катализируют гидролиз неорганического пирофосфата, превращая его в две молекулы фосфата и обеспечивая необходимые условия для таких реакций, как синтез белка, РНК и ДНК, делая таким образом эти ферменты незаменимыми для жизни [1]. Растворимые РРазы можно разделить на три структурно-независимых семейства (I, II, III). РРазы семейства II были открыты в 1998 г. [2, 3], и их до сих пор интенсивно исследуют [4, 5]. Эти ферменты существуют, как правило, в бактериях и археобактериях, включая патогены человека. Известно лишь несколько трехмерных структур высокого разрешения РРазе семейства II [6–8]. Канонические РРазы семейства II состоят из двух каталитических доменов, связанных гибким линкером между N-концевым ДНН- и C-концевым ДННА2-доменом и относятся к ДНН-гидролазам, названным так по консервативному мотиву Asp–His–His [9]. Фермен-

ты являются гомодимерами, в которых контакт осуществляется через ДНН-домены (за счет формирования объединенного β-листа). Около четверти ферментов РРаз семейства II содержит вставку из 250 остатков в ДНН-домене [5]. Вставка состоит из двух CBS-доменов (примерно по 60 аминокислотных остатков), названных в честь фермента цистатион-β-синтазы, в которой они были идентифицированы впервые, и DRTGG-домена (около 120 аминокислотных остатков). Эти РРазы обозначаются как CBS-РРазы. Известна кристаллическая структура регуляторной вставки из *Clostridium perfringens*, в которой пара CBS- и DRTGG-доменов образуют гомодимерную структуру. CBS-домены довольно широко распространены в природе и обнаружены, например, в 75 белках человека и восьми белках *Escherichia coli*. Мутации в CBS-доменах ферментов и мембранных каналов у человека связаны с рядом наследственных заболеваний, включая гомоцистинурию, пигментозный ретинит, синдром Бартера и остеопетроз [10]. CBS-домены всегда встречаются парами и образуют функциональный

псевдодимер (модуль Бейтмана), способный связывать различные лиганды (производные аденозина, гуанозина, ионы магния и другие) [11]. Кроме того, CBS-домены участвуют в образовании структур более высокого порядка, формируя из двух модулей Бейтмана гомодимерные структуры – CBS-модули, состоящие уже из четырех CBS-доменов. В качестве лигандов CBS-доменов в CBS-PPазах выступают различные фосфатные производные аденозина [12–14]. Аденозинмонофосфат (АМФ) и аденозиндифосфат (АДФ) ингибируют CBS-PPазы, тогда как аденозинтрифосфат (АТФ) и диаденозинтетрафосфат (Ар₄А) активируют его. Хотя каждая пара CBS-доменов содержит две потенциальные полости для связывания лигандов (CBS-модуль – четыре), регуляторная вставка CBS-PPазы связывает только две молекулы АМФ или одну молекулу Ар₄А через оба аденозиновых фрагмента [15].

Несмотря на физиологическую важность регуляции пирозинфосфатаз CBS-доменами, детали базовых молекулярных механизмов функционирования этих ферментов начали выяснять недавно [16, 17]. Для полного понимания путей передачи сигнала между их активным и регуляторным центрами необходима структурная информация на атомном уровне, которая, к сожалению, на данный момент недоступна из-за трудностей кристаллизации полноразмерных CBS-PPаз. Таким образом, их структурное исследование остается актуальной задачей. Метод малоуглового рентгеновского рассеяния (МУРР) был выбран в качестве основного исследовательского подхода, поскольку его можно использовать для изучения структуры биологических объектов непосредственно в растворе, т.е. в условиях, наиболее близких к физиологическим [18, 19]. Целью настоящей работы был структурный анализ CBS-PPаз из *D. hafniense* (dh-PPаза) с помощью МУРР для получения информации о четвертичной структуре и олигомерном состоянии белка при его взаимодействии с лигандами. Результаты работы могут стать шагом к пониманию принципиально важных функциональных особенностей CBS-PPаз.

МАТЕРИАЛЫ И МЕТОДЫ

Выделение и очистка CBS-PPазы. Dh-PPаза выделена из клеток *E. coli* штамма BL21, трансформированных вектором pET-42b (Novagen), несущим соответствующий ген, как описано ранее [20], при использовании буфера Mops-KOH концентрацией 0.1 М, pH 7.2, 2 мМ MgCl₂, 0.1 мМ CoCl₂ и 150 мМ KCl на стадии гель-фильтрации. Полученный белковый препарат был электрофоретически гомогенным.

Концентрацию белка определяли спектрофотометрически, используя значение $A_{280}^{0.1\%} = 0.478$, рассчитанное из аминокислотного состава dh-PPазы по программе ProtParam [21]. Все эксперименты проводили в буфере Mops-KOH концентрацией 0.1 М, pH 7.2, 2 мМ MgCl₂, 0.1 мМ CoCl₂ и 150 мМ KCl.

Эксперимент и анализ данных МУРР. Измерения проводили с помощью метода МУРР на станции EMBL-P12 BioSAXS накопительного кольца PETRAIII синхротрона DESY Европейской Лаборатории молекулярной биологии (EMBL) в Гамбурге, Германия. Станция оборудована роботизированной системой подачи и смены образца, 2D-детектором Pilatus 2M (DECTRIS, Switzerland). Интенсивность рассеяния регистрировали в диапазоне векторов $0.027 < s < 5.0 \text{ nm}^{-1}$, длина волны рентгеновского излучения $\lambda = 0.124 \text{ nm}$ [22]. Измерения проводили в том же буфере, в котором был приготовлен образец, при температуре 10°C и постоянном контроле радиационного повреждения образца. Диапазон концентраций растворов белков варьировался от 1 до 3 мг/мл. Первичную обработку полученных данных проводили по стандартным методикам с использованием программы PRIMUS [23].

Радиус инерции R_g рассчитывали в приближении Гинье, которое справедливо в области $sR_g < 1.3$:

$$I_{\text{exp}}(s) = I(0)\exp(-s^2 R_g^2/3), \quad (1)$$

где интенсивность рассеяния в нулевой угол $I(0)$ определяется из экспериментальных данных $I_{\text{exp}}(s)$ [24] и характеризует общее количество рассеивающей материи, пропорциональное квадрату молекулярной массы рассеивающей частицы.

Молекулярные массы рассчитывали по данным МУРР двумя методами: с помощью подхода Байеса ($\text{MM}_{\text{Bayesian}}$) [24], в котором молекулярные массы определяли, в том числе, по интенсивности рассеяния в нулевой угол $I(0)$ и на основе объема Порода V_p (MM_{Porod}) [25]. Последнее значение было определено с учетом того, что эмпирическое соотношение между V_p и молекулярной массой белка приблизительно равно 1.65 [26].

Функция парных расстояний $p(r)$ и максимальный размер рассеивающих частиц D_{max} рассчитаны с помощью программы GNOM [27]:

$$p(r) = \frac{1}{2\pi^2} \int_0^{\infty} sr I(s) \sin(sr) ds. \quad (2)$$

Формы низкого разрешения белков dh-PPase получены с помощью метода *ab initio* и программы DAMMIN [28], которая использует алгоритм имитации отжига для построения моделей, соот-

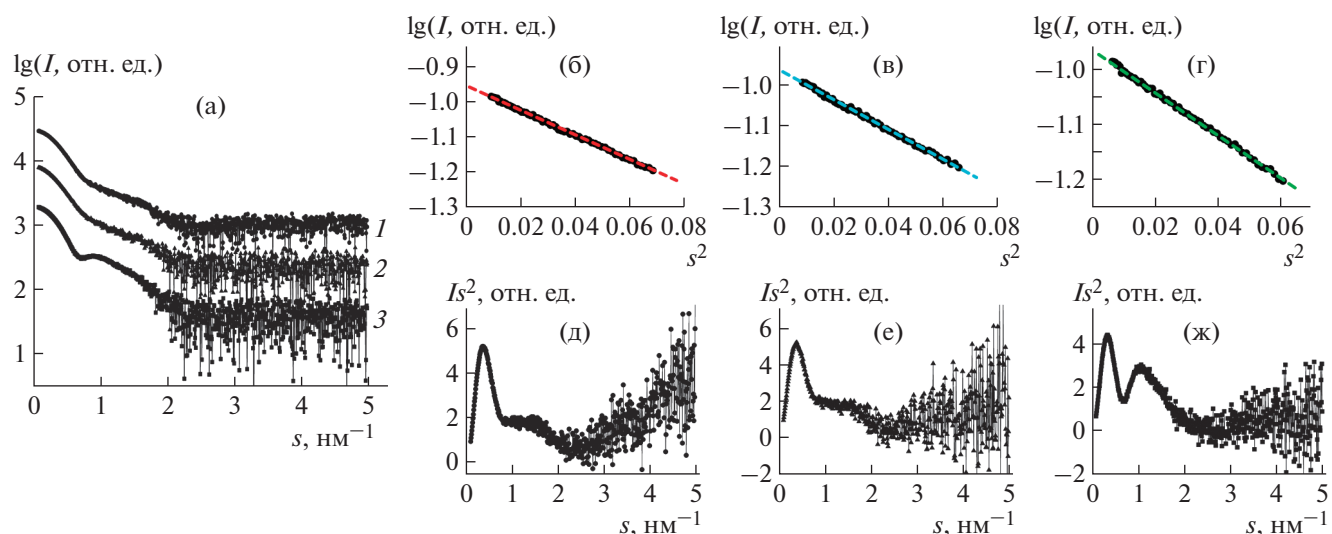


Рис. 1. Экспериментальные кривые рассеяния белка dh-PPазы (а): без лиганда (1), с АМФ (2), с Ар₄А (3). Графики Гинье (б, в, г) и Кратки (д, е, ж) для dh-PPазы без лиганда (б, д), с АМФ (в, е), с Ар₄А (г, ж).

ветствующих экспериментальным данным $I_{\text{exp}}(s)$, что минимизирует расхождение

$$\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I_{\text{exp}}(s_j) - cI_{\text{calc}}(s_j)}{\sigma(s_j)} \right]^2, \quad (3)$$

где N – число экспериментальных точек, c – коэффициент масштабирования, $I_{\text{calc}}(s_j)$ и $\sigma(s_j)$ – расчетная интенсивность в модели и экспериментальная ошибка интенсивностей соответственно.

Анализ неоднозначности полученных *ab initio*-моделей выполнен с помощью программы AMBIMETER [29]. Прогнозирование форм белков и их структурную классификацию проводили с помощью DATCLASS [26].

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Для исключения влияния на интерпретацию данных МУРР возможного межчастичного взаимодействия в растворе кривые малоуглового рассеяния dh-PPазы измеряли при нескольких концентрациях белка в диапазоне значений от 1 до 3 мг/мл. В указанном интервале концентрационной зависимости не наблюдалось. Поэтому с целью уменьшения экспериментальных ошибок для структурного анализа и вычисления макромолекулярных характеристик белка по данным МУРР использовали кривые, измеренные при концентрации 3 мг/мл (50 мкМ), т.е. кривые с минимальными экспериментальными шумами. Для изучения структурных перестроек при взаимодействии с лигандами в раствор исходной dh-PPазы вводили АМФ или Ар₄А с конечным содержанием лигандов 1000 и 100 мкМ соответ-

ственно. Полученные кривые малоуглового рассеяния представлены на рис. 1.

Как видно из рис. 1, профили экспериментальных кривых МУРР исходной dh-PPазы и кривых рассеяния этим ферментом при взаимодействии с лигандами АМФ и Ар₄А такие же, как в случае монодисперсных соединений – на них не наблюдается увеличения интенсивности рассеяния в самых малых углах. О монодисперсности образцов свидетельствуют также графики в координатах Гинье (рис. 1б–1г): для этих образцов приближение Гинье выполняется в достаточно большом интервале углов при $sR_g = 0.47$ – 1.29 для исходной dh-PPазы, 0.46 – 1.29 для dh-PPазы в комплексе с АМФ и 0.42 – 1.29 для dh-PPазы в комплексе с Ар₄А.

Графики Кратки (рис. 1д–1ж) для всех изучаемых dh-PPаз имеют характерный колоколообразный вид, что указывает на то, что все образцы в основном структурированы и имеют ограниченное число разупорядоченных, гибких областей (линкеров) [30]. Наибольшее количество разупорядоченных регионов в белке в соответствии с графиком Кратки наблюдается в исходной dh-PPазе – заметно большее увеличение $I s^2$ при увеличении вектора рассеяния s . Вероятно, присоединение лигандов в какой-то степени дополнительно структурирует фермент. Обращает на себя внимание график в координатах Кратки для комплекса dh-PPазы с Ар₄А. Резко выраженный вторичный максимум на этом графике может свидетельствовать о формировании компактной, но полой структуры. Это предположение будет рассмотрено ниже при обсуждении моделирования *ab initio* формы dh-PPазы в растворе. Кроме того,

Таблица 1. Макромолекулярные характеристики dh-PPазы с лигандами Ар₄А и АМФ и без них

Образец	R_g , нм	V_p , нм ³	D_{max} , нм	ММ _{Porod} , кДа	ММ _{Bayesian} , кДа
dh-PPаза	4.96 ± 0.45	410 ± 40	18	248 ± 20	242
dh-PPаза с АМФ	5.06 ± 0.23	388 ± 40	18.5	235 ± 20	243
dh-PPаза с Ар ₄ А	5.22 ± 0.03	403 ± 40	17	244 ± 20	242

именно для комплекса dh-PPазы с Ар₄А наблюдается наиболее значительное изменение профиля кривой МУРР, что говорит о существенных конформационных изменениях фермента при связывании Ар₄А (рис. 1а, кривая 3).

В табл. 1 представлены макромолекулярные характеристики (инварианты рассеяния) исследуемых CBS-PPаз. Молекулярные массы, определенные по экспериментальной кривой рассеяния, исходя из оценки объема Порода [25] и обобщенного подхода Байеса [24], равны в среднем 242 кДа, что соответствует молекулярной массе тетрамера. Эта величина практически совпадает с теоретическим значением молекулярной массы этого белка, рассчитанным по аминокислотной последовательности мономера (UniProt: B8FP42), умноженной на четыре, т.е. равной 241 кДа.

Радиусы инерции R_g как свободной dh-PPазы, так и в комплексах с АМФ и Ар₄А оказались близкими по величине в пределах ошибки определения. В то же время наблюдается тенденция к увеличению R_g при присоединении лигандов, причем R_g образца dh-PPазы в комплексе с Ар₄А, т.е. в случае присоединения тетрафосфата, несколько больше, чем при взаимодействии с монофосфатом (АМФ). Возможно, это объясняется стерическим фактором, поскольку размер Ар₄А больше размера АМФ, что приводит к связыванию только одной молекулы тетрафосфата на модуль Бейтмана, в то время как туда могут поместиться две молекулы АМФ [15].

Породовский объем V_p в пределах ошибки измерения одинаков для всех образцов. Этот параметр характеризует суммарное рассеяние образцом, и вклад небольших молекул лигандов практически не сказывается на его значении. Основной вывод этой части исследования заключается в констатации факта образования устойчивых тетрамеров dh-PPаз, в том числе в случае взаимодействия этого фермента с разными лигандами.

Для определения *ab initio* форм низкого разрешения dh-PPазы с помощью программы DAMMIN [28] были рассчитаны функции парных расстояний $p(r)$ (рис. 2б–2г). Чтобы избежать ошибочной интерпретации, функции парных расстояний $p(r)$ были рассчитаны по укороченным кривым малоуглового рассеяния в интервале векторов $0.027 < s < 1.6 \text{ нм}^{-1}$. Выбор такого интервала обусловлен тем, что использование полных кривых рассея-

ния может привести к появлению в восстановленных формах ложных мелких деталей. Метод восстановления *ab initio*, основанный на имитации отжига, позволяет получить форму рассеивающего объекта с разрешением порядка 2 нм. Для этого используют начальную часть кривой малоуглового рассеяния, длина которой определяется ее информационной составляющей: число шенноновских каналов $N_s = s_{max} D_{max} / 2\pi$ должно быть порядка десяти [31]. В данном случае для восстановления структуры низкого разрешения dh-PPазы использовали девять шенноновских каналов, для комплекса с АМФ — десять, а с Ар₄А — девять. Поскольку исследуемая dh-PPаза представляет собой тетрамер, образованный двумя гомодимерами, ее структуру определяли с использованием симметрии $P2$.

Восстановленные таким образом формы представлены на рис. 2д–2ж. О хорошем качестве и надежности восстановления свидетельствуют значения $\chi^2 = 1.21$ для исходной dh-PPазы, 1.08 и 0.94 для комплексов с АМФ и Ар₄А соответственно. Dh-PPаза при связывании Ар₄А формирует полую структуру, представленную восстановленной формой этого белка (рис. 2ж). Возможность формирования полой структуры dh-PPазы в комплексе с Ар₄А была предсказана при анализе графика Кратки, а также с помощью программы DATCLASS [25], тогда как для dh-PPазы без лигандов или с АМФ программой DATCLASS были предсказаны просто компактные структуры, показанные на рис. 2д, 2е.

Поскольку решение обратной задачи восстановления трехмерных форм по одномерной кривой рассеяния неоднозначно, важно подчеркнуть, что анализ неоднозначности полученных *ab initio*-моделей, проведенный с помощью программы AMBIMETER [29], показал, что индекс неоднозначности найденных моделей меньше 1.5. Этот результат указывает на практически уникальные модели белка.

Таким образом, при использовании функции парных расстояний, симметрии $P2$ и метода *ab initio* восстановления формы, реализованного в программе DAMMIN [28], были получены формы низкого разрешения белка исходной dh-PPазы и в комплексе с различными лигандами. Наиболее заметные конформационные изменения в четвертичной структуре белка происходят при

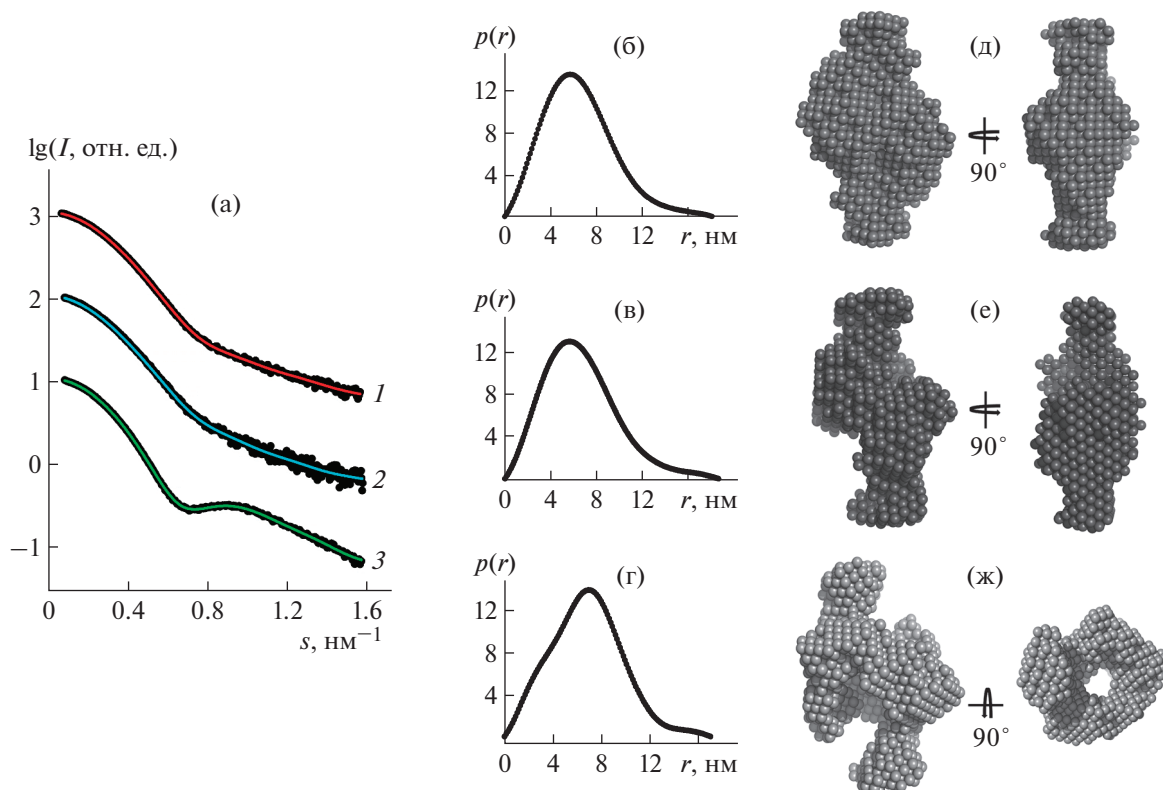


Рис. 2. Экспериментальные кривые рассеяния (точки) и модельные кривые (сплошные линии), рассчитанные с помощью программы DAMMIN для форм низкого разрешения CBS-пирофосфатаз (а): dh-PPазы без лиганда (1), с АМФ (2), с Ар₄А (3). Кривые смещены по вертикали на одну логарифмическую единицу для лучшей визуализации. Функции парных расстояний (б, в, г) и восстановленные формы белка (д, е, ж) dh-PPазы без лиганда (б, д), с АМФ (в, е) и Ар₄А (г, ж).

связывании Ар₄А. Можно предположить, что такие изменения вызваны природой Ар₄А, поскольку этот лиганд связывается с CBS-доменами своими двумя аденозиновыми фрагментами, формируя более открытую конформацию регуляторной вставки по сравнению с АМФ [15]. Лиганд Ар₄А, по-видимому, расталкивает другие фрагменты белка, переводя его в активное состояние. Это заключение отчасти подтверждается тем, что Ар₄А активирует CBS-PPазу, тогда как АМФ ингибирует ее [12, 13]. То есть полученную полую структуру фермента в комплексе с Ар₄А можно рассматривать в качестве активной конформации dh-PPаз, а в комплексе с АМФ — в качестве малоактивной. В целом взаимодействия субъединиц должны быть надлежащим образом сбалансированы, чтобы обеспечить передачу конформационных изменений между ними [17].

Участие регуляторных CBS-доменов в образовании не только димера, но и тетрамера соотносится с информацией о том, что канонические PPазы семейства II (содержащие только каталитические домены) являются гомодимерами [9].

ЗАКЛЮЧЕНИЕ

Полученная методом МУРР информация об олигомерной структуре dh-PPазы (тетрамер, образованный двумя гомодимерами) подтверждает многоуровневый механизм регуляции CBS-PPазы. Результаты проведенных структурных исследований могут служить базисом для понимания механизма передачи информации между регуляторными и каталитическими центрами.

Авторы выражают благодарность профессору Р. Лахти (Университет Турку, Турку, Финляндия) за предоставленную плазмиду с геном dh-PPазы.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 18-34-00918) и Министерства науки и высшего образования РФ в рамках Государственного задания ФНИЦ «Кристаллография и фотоника» РАН.

СПИСОК ЛИТЕРАТУРЫ

1. *Heinonen J.* Biological Role of Inorganic Pyrophosphate. М.: Kluwer Academic Publishers, 2001. 250 p.
2. *Shintani T., Uchiumi T., Yonezawa T. et al.* // FEBS Lett. 1998. V. 439. P. 263.

3. *Young T., Kuhn N., Wadeson A. et al.* // *Microbiology*. 1998. V. 144. P. 2563.
4. *Anashkin V.A., Aksenova V.A., Salminen A. et al.* // *BBRC*. 2015. V. 517. P. 266.
5. *Baykov A.A., Anashkin V.A., Salminen A. et al.* // *FEBS Lett.* 2017. V. 591. P. 3225.
6. *Merckel M., Fabrichnyy I., Salminen A. et al.* // *Structure*. 2001. V. 94. P. 289.
7. *Ahn S., Milner A., Fütterer K. et al.* // *J. Mol. Biol.* 2001. V. 313. P. 797.
8. *Rantanen M., Lehtiö L., Rajagopal L. et al.* // *Acta Cryst. D*. 2007. V. 63. P. 738.
9. *Aravind L., Koonin E.* // *Trends Cell Biol.* 1998. V. 23. P. 17.
10. *Ignoul S., Eggermont J.* // *Am. J. Physiol. Cell Physiol.* 2005. V. 289. P. 1369.
11. *Ereno-Orbea J., Oyenarte I., Martinez-Cruz L.A.* // *Arch. Biochem. Biophys.* 2013. V. 540. P. 70.
12. *Jamsen J., Tuominen H., Salminen A. et al.* // *Biochem. J.* 2007. V. 408. P. 327.
13. *Anashkin V.A., Salminen A., Tuominen H.K. et al.* // *J. Biol. Chem.* 2015. V. 290. P. 27594.
14. *Salminen A., Anashkin V.A., Lahti M. et al.* // *J. Biol. Chem.* 2014. V. 289. P. 22865.
15. *Tuominen H., Salminen A., Oksanen E. et al.* // *J. Mol. Biol.* 2010. V. 398. P. 400.
16. *Anashkin V.A., Salminen A., Osipova E. et al.* // *ACS Omega*. 2019. V.4. P. 15549.
17. *Anashkin V.A., Salminen A., Vorobjeva N.N. et al.* // *Biochem. J.* 2016. V. 473. P. 2097.
18. *Svergun D.I., Koch M.H.J., Timmins P.A. et al.* *Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules*. Oxford: Oxford University Press, 2013. 358 p.
19. *Korasick D.A., Tanner J.J.* // *Protein Sci.* 2018. V. 27. P. 814.
20. *Parfenyev A., Salminen A., Halonen P. et al.* // *J. Biol. Chem.* 2001. V. 276. P. 24511.
21. *Gasteiger E., Hoogland C., Gattiker A. et al.* *The Proteomics Protocols Handbook*. N.J.: Humana Press, 2005. 607 p.
22. *Jeffries C.M., Graewert M.A., Svergun D.I. et al.* // *J. Synchr. Rad.* 2015. V. 22. P. 273.
23. *Konarev P.V., Volkov V.V., Sokolova A.V. et al.* // *J. Appl. Cryst.* 2003. V. 36. P. 1277.
24. *Hajizadeh N.R., Franke D., Jeffries C.M. et al.* // *Sci. Rep.* 2018. V. 8. P. 7204.
25. *Porod G., Glatter O., Kratky O. et al.* *Small angle X-ray Scattering*. London: Academic Press, 1982. 51 p.
26. *Franke D., Petoukhov M.V., Konarev P.V. et al.* // *J. Appl. Cryst.* 2017. V. 50. P. 1212.
27. *Svergun D.I.* // *J. Appl. Cryst.* 1992. V. 25. P. 495.
28. *Svergun D.I.* // *Biophys. J.* 1999. V. 76. P. 2879.
29. *Petoukhov M.V., Svergun D.I.* // *Acta Cryst. D*. 2015. V. 71. P. 1051.
30. *Jacques D.A., Guss J.M., Svergun D.I. et al.* // *Acta Cryst. D*. 2012. V. 68. P. 620.
31. *Konarev P.V., Svergun D.I.* // *IUCr J.* 2015. V. 2. P. 352.