

Анализ структуры белковых молекул с посттрансляционными модификациями при онкопатологии

Тихонов Д.А.^{1,2}, Куликова Л.И.^{1,2}, Копылов А.Т.³, Руднев В.Р.², Кайшева А.Л.³

¹Институт математических проблем биологии РАН – филиал ИПМ РАН
им. М.В. Келдыша, Пуццино, Московская область, Россия

²Институт теоретической и экспериментальной биофизики РАН,
Пуццино, Московская область, Россия

³Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича,
Москва, Россия

likulikova@mail.ru

Авторами предложен новый биоинформационный подход исследования идентифицированных аминокислотных посттрансляционных модификаций в белковых структурах с помощью пространственного описания их ближайшего окружения. Было впервые установлено, что белковые фрагменты, последовательности целевых пептидов, являются составной частью спиральных пар. Были определены координаты этих структур, проведен анализ основных характеристик содержащих выявленные модификации спиральных пар: межспиральные расстояния, торсионный угол между осями спиралей, количество аминокислот между спиральями, длина спиралей, площадь и периметр пересечения проекций спиралей. Показано, что посттрансляционные модификации в белковых структурах экспонированы на водное окружение – растворитель. Проведен анализ распределения белковых молекул, содержащих различные типы модификаций, в зависимости от площади доступной растворителю поверхности до и после модификации. Установлено, что площадь доступной растворителю поверхности модифицированного аминокислотного остатка всегда увеличивается по сравнению с интактным немодифицированным аминокислотным остатком. Разработанный подход дает возможность по-новому оценить вероятностные изменения белка, в котором обнаружены посттрансляционные модификации, и помогут делать прогнозы биологической активности данного белка.

Ключевые слова: посттрансляционные модификации, спиральная пара, структурные мотивы белковых молекул, основные характеристики структур.

Analysis of Protein Molecule Structure with Post-Translational Modifications in Oncopathology

Tikhonov D.A.^{1,2}, Kulikova L.I.^{1,2}, Kopylov A.T.³, Rudnev V.R.², Kaysheva A.L.³

¹Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics RAS,
Pushchino, Russia

²Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Pushchino, Russia

³V.N. Orekhovich Institute of Biomedical Chemistry, Moscow, Russia

The authors have proposed a new bioinformatic approach to studying the identified amino acid post-translational modifications in protein structures using a spatial description of their immediate environment. It was first established that protein fragments, sequences of target peptides, are an integral part of helical pairs. The coordinates of these structures were determined, the analysis of the main characteristics containing the identified modifications of helical pairs was carried out: inter-helix distances, torsion angle between the axes of the helices, the number of amino acids between the helices, the length of the helices, the area and perimeter of the intersection of the projections of the helices. It is shown that post-translational modifications in protein structures are exposed to an aqueous environment – a solvent. The analysis of the distribution of protein molecules containing various types of modifications, depending on the area of the surface accessible to the solvent before and after modification. It was found that the area accessible to the solvent of the surface of the modified amino acid residue always increases in comparison with the intact unmodified amino acid residue. The developed approach makes it possible to re-evaluate the probabilistic changes in the

protein in which post-translational modifications are found, and will help to make predictions of the biological activity of this protein.

Key words: post-translational modifications; helical pair, protein structural motifs.

1. Введение

Решающая роль в осуществлении биологической активности белка отводится исследователями трехмерной структуре и особенностям фолдинга [1], однако модификация аминокислотных остатков вносит существенный вклад в структурное и функциональное разнообразие белков [2, 3]. Посттрансляционные модификации (ПТМ) белков влияют на ферментативную активность, локализацию белка, белок-белковые взаимодействия, регуляцию сигнальных каскадов, репарацию ДНК и деление клеток [4]. В связи с этим не прекращаются попытки выяснить роль ПТМ белка в развитии различных патологий. Понимание роли различных ПТМ, управляющих метаболизмом рака, является сложной задачей из-за сложной и гетерогенной природы опухолей. Таким образом, необходимы дальнейшие исследования и интеграция данных ПТМ для разработки более точных стратегий, которые могут предложить новые подходы к нацеливанию на метаболизм рака. Работа выполнена на базе проведенного сравнительного масс-спектрометрического анализа [5–7] белкового состава с учетом посттрансляционных модификаций образцов плазмы крови здоровых добровольцев и образцов больных колоректальным раком для трех типов биологических значимых модификаций – фосфорилирование, ацетилирование, убиквитинирование. Предложен биоинформационный подход пространственного описания окружения идентифицированных аминокислотных посттрансляционных модификаций в белковых структурах для определения структурного мотива [8–10], содержащего модификацию, и проведения конформационного анализа.

2. Методы

2.1. Идентификация белков

Белки, содержащие посттрансляционные модификации, были идентифицированы с помощью тандемного масс-спектрометрического анализа белкового состава образцов плазмы крови двух серий сравнения – больных CRC ($n = 28$) и здоровых добровольцев ($n = 41$) [11]. Были выявлены 6 белков, содержащих 12 пептидов с модифицированными аминокислотными остатками, для которых аннотированы интактные 3D-структуры в базе данных Protein Data Base [12].

2.2. Анализ посттрансляционных модификаций белков

Для настоящего исследования из Банка белковых данных (PDB) были отобраны все белки, содержащие триптические пептиды определенного типа модификации. Для каждого такого пептида была создана своя выборка для определения конформационного шаблона. Далее из полученной базы данных были отобраны все мотивы, содержащие исследуемые триптические пептиды. Для распознавания и отбора структурных мотивов использовался метод, разработанный нами ранее [13–15]. Вторичная структура белков определялась методом DSSP Кабша и Сандера [16]. С помощью этой же программы определялись доступные поверхности контакта с растворителем. Определения важных характеристик структурных мотивов описаны в опубликованных ранее работах [13–15]. Визуальный анализ структур проводился с помощью программы молекулярной графики RasMol [17].

3. Результаты

Характеризация пространственного окружения выявленных модифицированных аминокислотных остатков в белковых глобулах при колоректальном раке представляет интересную научную задачу для более глубокого понимания механизмов развития онкопатологии. Результатом решения этой задачи является создание дополнительного инструмента валидации математически предсказанных ПТМ белков/пептидов путем сравнительного анализа изменений массово-зарядных характеристик триптических пептидов и их фрагментов с учетом модификаций и соответствующих пептидов без модификаций.

В нашем исследовании выполнен анализ пространственного окружения модифицированных аминокислотных остатков в белках, выявленных в образцах крови больных колоректальным раком: Vitamin D - binding protein (VTDB), Alpha-2-macroglobulin (A2MG), Complement C4-A (CO4A), X-ray repair cross-complementing protein 6 (XRCC6), Plasmaprotease C1 inhibitor (IC1), Serumalbumin (ALBU).

Предложен новый подход исследования идентифицированных аминокислотных посттрансляционных модификаций в белковых структурах с помощью пространственного описания их ближайшего окружения: определения структурных мотивов, содержащих выявленные модификации, и анализа основных характеристик мотивов. Это дает возможность по-новому оценить

вероятностные изменения белка, в котором обнаружены ПТМ, а значит, и сделать прогнозы биологической активности данного белка. В данной работе впервые были выявлены структурные мотивы белков, содержащие пострасляционные модификации, в образцах крови больных колоректальным раком, и проведен их конформационный анализ. Для этого из Банка белковых данных отобраны все белковые молекулы, содержащие изучаемые триптические пептиды, охарактеризованы конформационные шаблоны для каждого пептида с ПТМ. Установлены координаты начала и конца последовательностей в структурах молекул белка. Определены образуемые ими типы структурных мотивов. Показано, что, как правило, триптические пептиды являются составной частью спиральных пар: $\alpha\alpha$ -уголков, L-структур, V-структур и т. д.

Проведено исследование отобранных спиральных пар: для каждой такой структуры были вычислены межспиральные расстояния (минимальное r и межплоскостное d), углы между осями спиралей (торсионный θ и двумерный ϕ), площадь (S) и периметр (P) пересечения проекций спиралей, длина перетяжки (N_p) между спиральями, длины спиралей. Также в работе проанализированы доступные для молекул воды поверхности модифицированного участка исследуемых триптических пептидов.

Мы показали, что все идентифицированные аминокислотные остатки с ПТМ локализованы в спиральных парах размерами от 22 до 66 аминокислотных остатков. При этом мы наблюдаем, что идентифицированные аминокислотные модификации локализованы в одной из спиралей либо на конце спирали в непосредственной близости к нерегулярному участку в гидрофильном кластере и не встречаются в гидрофобном кластере. Аминокислотные остатки с ПТМ в белковых структурах в составе выборок конформационных шаблонов экспонированы в водное окружение (растворитель). Мы показали, что площадь поверхности модифицированного аминокислотного остатка, доступная для растворителя, умеренно превосходит таковую для интактных немодифицированных структур, что говорит о небольшом изменении локальной конформации белка – "дыхание молекулы". Такие конформационные изменения состояния являются вероятными (карта разрешенных состояний Рамачандрана) и, вероятно, не приводят к расправленной глобуле и гибели белка.

Мы предполагаем, что модификации белков, которые мы исследовали, участвуют в регуляции функциональной активности белка, поскольку зачастую располагаются в пространстве вблизи сайтов связывания с молекулами-партнерами.

Проведенные исследования актуальны и важны, поскольку полученные знания могут быть полезными в понимании вклада модификации аминокислотных остатков в регуляцию

биологической активности белков при колоректальных патологиях.

4. Благодарности

Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований, проект № 18-07-01031-а.

4. Список литературы

1. Díaz-Villanueva J.F., Díaz-Molina R., García-González V. Protein Folding and Mechanisms of Proteostasis. *Int. J. Mol. Sci.* 2015. V. 16. P. 17193–17230. doi: [10.3390/ijms160817193](https://doi.org/10.3390/ijms160817193).
2. Karve T.M., Cheema A.K. Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *Journal of Amino Acids.* 2011. 13. doi: [10.4061/2011/207691](https://doi.org/10.4061/2011/207691).
3. Karve T., Cheema A. Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *J. Amino Acids.* 2011. Article ID 207691. doi: [10.4061/2011/207691](https://doi.org/10.4061/2011/207691).
4. *Post-Translational Modifications in Health and Disease.* Ed. Vidal C.J. New York, USA: Springer New York, 2011. doi: [10.1007/978-1-4419-6382-6](https://doi.org/10.1007/978-1-4419-6382-6).
5. Sundararajan N., Mao D., Chan S., Koo T.-W., Su X., Sun L., Zhang J., Sung K., Yamakawa M., Gafken P.R. et al. Ultrasensitive Detection and Characterization of Posttranslational Modifications Using Surface-Enhanced Raman Spectroscopy. *Anal. Chem.* 2006. V. 78. P. 3543–3550. doi: [10.1021/ac051525i](https://doi.org/10.1021/ac051525i).
6. Li X., Kapoor T.M. Approach to Profile Proteins That Recognize Post-Translationally Modified Histone "Tails." *J. Am. Chem. Soc.* 2010. V. 132. P. 2504–2505. doi: [10.1021/ja909741q](https://doi.org/10.1021/ja909741q).
7. Khidekel N., Hsieh-Wilson L.C. A 'Molecular Switchboard'—Covalent Modifications to Proteins and Their Impact on Transcription. *Org. Biomol. Chem.* 2004. V. 2. P. 1–7. doi: [10.1039/B312466E](https://doi.org/10.1039/B312466E).
8. Efimov A.V. Standard Structures in Proteins. *Prog. Biophys. Mol. Biol.* 1993. V. 60. P. 201–239. doi: [10.1016/0079-6107\(93\)90015-C](https://doi.org/10.1016/0079-6107(93)90015-C).
9. Brazhnikov E., Efimov A. Structure of $\alpha\alpha$ -Hairpins with Short Connections in Globular Proteins. *Mol. Biol.* 2001. V. 35. P. 89–97. doi: [10.1023/A:1004859003221](https://doi.org/10.1023/A:1004859003221).
10. Efimov A.V. L-shaped structure from two α -helices with a proline residue between them. *Mol. Biol. (Mosk.)*. 1992. V. 26. P. 1370–1376.
11. Tikhonov D., Kulikova L., Kopylov A., Malsagova K., Stepanov A., Rudnev V., Kaysheva A. Super Secondary Structures of Proteins with Post-Translational Modifications in

- Colon Cancer. *Molecules*. 2020. V. 25. P. 3144. doi: [10.3390/molecules25143144](https://doi.org/10.3390/molecules25143144).
12. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. The Protein Data Bank. *Nucl. Acids Res.* 2000. V. 28. P. 235–242. doi: [10.1093/nar/28.1.235](https://doi.org/10.1093/nar/28.1.235).
 13. Tikhonov D.A., Kulikova L.I., Efimov A.V. The Study of Interhelical Angles in the Structural Motifs Formed by Two Helices. *Mathematical Biology and Bioinformatics*. 2019. V. 14. P. 1–17. doi: [10.17537/2019.14.t1](https://doi.org/10.17537/2019.14.t1).
 14. Tikhonov D.A., Kulikova L.I., Efimov A.V. Analysis of torsion angles between helical axes in pairs of helices in protein molecules. *Mathematical Biology and Bioinformatics*. 2018. V. 13. P. 17–28. doi: [10.17537/2018.13.t17](https://doi.org/10.17537/2018.13.t17).
 15. Tikhonov D.A., Kulikova L.I., Efimov A.V. Statistical analysis of internal distances of helical pairs in protein molecules. *Mathematical Biology and Bioinformatics*. 2019. V. 14. P. 18–36. doi: [10.17537/2019.14.t18](https://doi.org/10.17537/2019.14.t18).
 16. Kabsch W., Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983. V. 22. P. 2577–2637. doi: [10.1002/bip.360221211](https://doi.org/10.1002/bip.360221211).
 17. Sayle R.A., Milner-White E.J. RASMOL: Biomolecular graphics for all. *Trends Biochem. Sci.* 1995. V. 20. P. 374–376. doi: [10.1016/s0968-0004\(00\)89080-5](https://doi.org/10.1016/s0968-0004(00)89080-5).