

**О ВКЛАДЕ СПЕЦИФИЧЕСКИХ ЧЕРЕДОВАНИЙ ГИДРОФОБНЫХ И
ГИДРОФИЛЬНЫХ АМИНОКИСЛОТНЫХ ОСТАТКОВ В
ФОРМИРОВАНИЕ ЭЛЕМЕНТОВ ВТОРИЧНОЙ СТРУКТУРЫ
БЕЛКОВ**

Хрусталёв В.В., Побойнев В.В., Хрусталёва Т.А.*

*Белорусский государственный медицинский университет, кафедра общей
химии*

**ГНУ «Институт физиологии НАН Беларуси»
г. Минск*

Ключевые слова: альфа-спираль, бета-тяж, вторичная структура.

Резюме: оригинальная вероятностная шкала, основанная на чередовании гидрофильных и гидрофобных аминокислотных остатков в составе пентапептидов, в большей степени подходит для выявления альфа-спиралей, чем традиционная вероятностная шкала для каждого из двадцати аминокислотных остатков; однако для выявления бета-тяжей необходимы обе шкалы.

Resume: the original probability scale based on combinations of hydrophilic and hydrophobic amino acid residues in pentapeptides is more accurate in finding alpha helices than the traditional amino acid propensity scale, while for accurate finding of beta strands the both scales are necessary.

Актуальность. Вторичная структура белка формируется под воздействием как внешних, так и внутренних факторов. К внутренним факторам относятся ближние и дальние взаимодействия между аминокислотными остатками одной и той же полипептидной цепи. Ближние взаимодействия можно предсказать с высокой долей вероятности, в отличие от дальних. Действительно, известны аминокислотные остатки, которые достоверно чаще образуют альфа-спирали (аланин, глутамин, глутаминовая кислота, аргинин, лейцин, метионин, лизин) [2], и те остатки, которые чаще образуют бета-тяжи (валин, изолейцин, цистеин, фенилаланин, тирозин, триптофан, треонин) [2]. Некоторые аминокислотные остатки чаще находятся за пределами элементов вторичной структуры (пролин, глицин, аспарагиновая кислота, аспарагин, серин, гистидин) [2]. Однако для большинства из аминокислотных остатков перечисленные выше предпочтения нельзя считать достаточно резко выраженными. Для формирования того или иного элемента вторичной структуры соответствующие аминокислоты должны находиться в непосредственной близости друг от друга – образовывать кластер. Именно такими представлениями пользовались создатели первых методов предсказания вторичной структуры по аминокислотной последовательности [2]. Как выяснилось, с помощью такой логики можно объяснить формирование лишь некоторой части альфа-спиралей и бета-тяжей.

Определённое чередование гидрофильных и гидрофобных аминокислотных остатков также играет важную роль в формировании альфа-спиралей и бета-тяжей [5]. Первые методы предсказания вторичной структуры, основанные на выявлении таких чередований, были способны выявить лишь некоторую часть структурных элементов [5]. Тем не менее, использование подобных подходов привело к пониманию того, что «альфа-спиральные» аминокислотные остатки могут чередоваться таким образом, что становятся более склонными к формированию бета-тяжей, а «бета-структурные» остатки могут попадать в комбинации, которые, наоборот, характерны для альфа-спиралей.

Влияние дальних взаимодействий на вторичную структуру заключается в том, что фрагменты полипептидной цепи, далёкие друг от друга по расположению в первичной последовательности, могут сближаться и индуцировать формирование элементов вторичной структуры, не получивших бы возможности сформироваться в случае отсутствия таких

контактов. Влияние дальних взаимодействий нивелируется при работе с короткими пептидами, соответствующими той или иной области полноразмерного белка. Для выделения структурно устойчивых пептидов необходим алгоритм, учитывающий как особенности самих аминокислотных остатков, так и характер их чередования. Кроме того, большой теоретический интерес представляет расчёт процентного соотношения вклада ближних и дальних взаимодействий в формирование элементов вторичной структуры.

Цель настоящего исследования: оценка вклада особенностей аминокислотного состава и специфического чередования гидрофильных и гидрофобных остатков в формирование альфа-спиралей и бета-тяжей в белках.

Задачи исследования включали: 1. Создание двух шкал для оценки вероятности формирования элементов вторичной структуры на основании анализа выборки негомологичных белков бактерий с различной GC-насыщенностью их геномов; 2. Обработка независимой выборки негомологичных белков человека и хордовых животных с помощью алгоритма, основанного на полученных вероятностных шкалах.

Материалы и методы. Материалом для создания алгоритма PentaFOLD (<http://chemres.bsmu.by>) послужили три выборки трёхмерных структур бактериальных белков из международной базы данных PDB (www.pdb.org). Максимальный процент сходства аминокислотных последовательностей этих белков друг с другом равен 25%. Общая численность белков составила 542. В первую выборку вошли белки GC-бедных бактерий, протеомы которых особенно обогащены изолейцином, аспарагином и лизином, в меньшей степени – фенилаланином, тирозином и метионином [4]. Вторая выборка включала белки GC-богатых бактерий, обогащенные аланином и, в меньшей степени, глицином, аргинином и пролином [4]. Третья выборка содержала белки бактерий со средней GC-насыщенностью [4].

На основании анализа аминокислотного состава альфа-спиралей, бета-тяжей и участков полипептидной цепи, находящихся в неструктурированном состоянии, была получена первая вероятностная шкала, являющаяся по сути обновлённым вариантом шкалы Chou и Fasman [2]. В алгоритме PentaFOLD вероятность включения аминокислотного остатка в тот или иной элемент вторичной структуры рассчитывается как среднее значение для пентапептида, в центре которого находится рассматриваемый остаток.

Чередования гидрофобных и гидрофильных аминокислотных остатков изучали во фрагментах длиной в пять аминокислотных остатков (в пентапептидах). Вторая вероятностная шкала PentaFOLD основана на частотах использования 32 типов пентапептидов (остатки в них разделены только на две категории: гидрофобные и гидрофильные, таким образом $2^5 = 32$ возможных вариантов пентапептидов) в альфа-спиралях, бета-тяжах и неструктурированных фрагментах.

Независимая выборка состояла из 50 трехмерных структур белков человека и хордовых животных, процент сходства между первичными последовательностями которых не превышал 25%. В составе каждого из структурных элементов нами рассчитан процент аминокислотных остатков, обладающих 1) наибольшей вероятностью включения в них исключительно по «аминокислотной» шкале, 2) наибольшей вероятностью включения исключительно по «пентапептидной» шкале, 3) наибольшими вероятностями включения в данный элемент вторичной структуры по обеим шкалам, 4) не обладающие наибольшей вероятностью включения ни по одной из двух шкал. Сравнение между полученными величинами проводили с помощью t-теста.

Результаты и их обсуждение. Результаты эксперимента *in silico* показали, что роль чередования гидрофобных и гидрофильных остатков в формировании альфа-спиралей является достоверно более важной, чем роль наличия кластеров аминокислот с высоким альфа-спиральным потенциалом. Действительно, 27,84% аминокислотных остатков в альфа-спиралях находятся в центре пентапептидов с чередованием остатков, характерным для альфа-спиралей (Рисунок 1). Только 21,20% аминокислот из альфа-спиралей находятся в центре кластера, состоящего из пяти остатков со средним альфа-спиральным потенциалом, превышающим таковой для бета-тяжей и неструктурированного состояния. Ещё более странным кажется тот факт, что процент остатков, находящихся среди альфа-спиральных аминокислот в альфа-спиральных комбинациях, составил всего 22,40%. Это говорит о том, что для образования альфа-спирали вполне достаточно и одного из двух исследуемых в этой работе факторов. С помощью хотя бы одной из двух вероятностных шкал можно объяснить включение только 71,44% остатков в состав альфа-спиралей. Более четверти остатков включаются в альфа-спирали без участия причин, связанных с их природой или с природой соседних аминокислот. Действительно, между альфа-спиралями внутри одного и того же белка часто формируются контакты, стабилизирующие структуру одной из них. Яркий пример таких контактов – большой прионный белок человека, в котором третья альфа-спираль стабилизируется с помощью контактирующей с ней второй альфа-спирали [1].

Среди аминокислотных остатков, формирующих бета-тяжи, достоверно чаще встречаются те, которые обладают высоким бета-потенциалом и находятся в центре бета-структурных пентапептидов (33,27%). То есть, для формирования бета-тяжа предпочтительно выполнение обоих требований: такие остатки как валин или изолейцин должны находиться или в гидрофобных пентапептидах, или в пентапептидах с чередованием гидрофобных и гидрофильных остатков через один, а не через два, как это происходит в альфа-спиралях [3]. В данном случае речь, по всей видимости, идёт об «активных» бета-тяжах, которые находят в близлежащем пространстве партнеров для образования параллельной или

антипараллельной бета-структуры. В состав «пассивных» бета-тяжей, судя по нашим данным, чаще должны входить фрагменты с бета-структурными аминокислотами в небетаструктурных комбинациях (которых всего 23,85%), чем с бета-структурным чередованием небетаструктурных остатков (их 18,87%). Включение в бета-тяжи невозможно объяснить ни по одному из методов для 24,00% остатков (Рисунок 1).

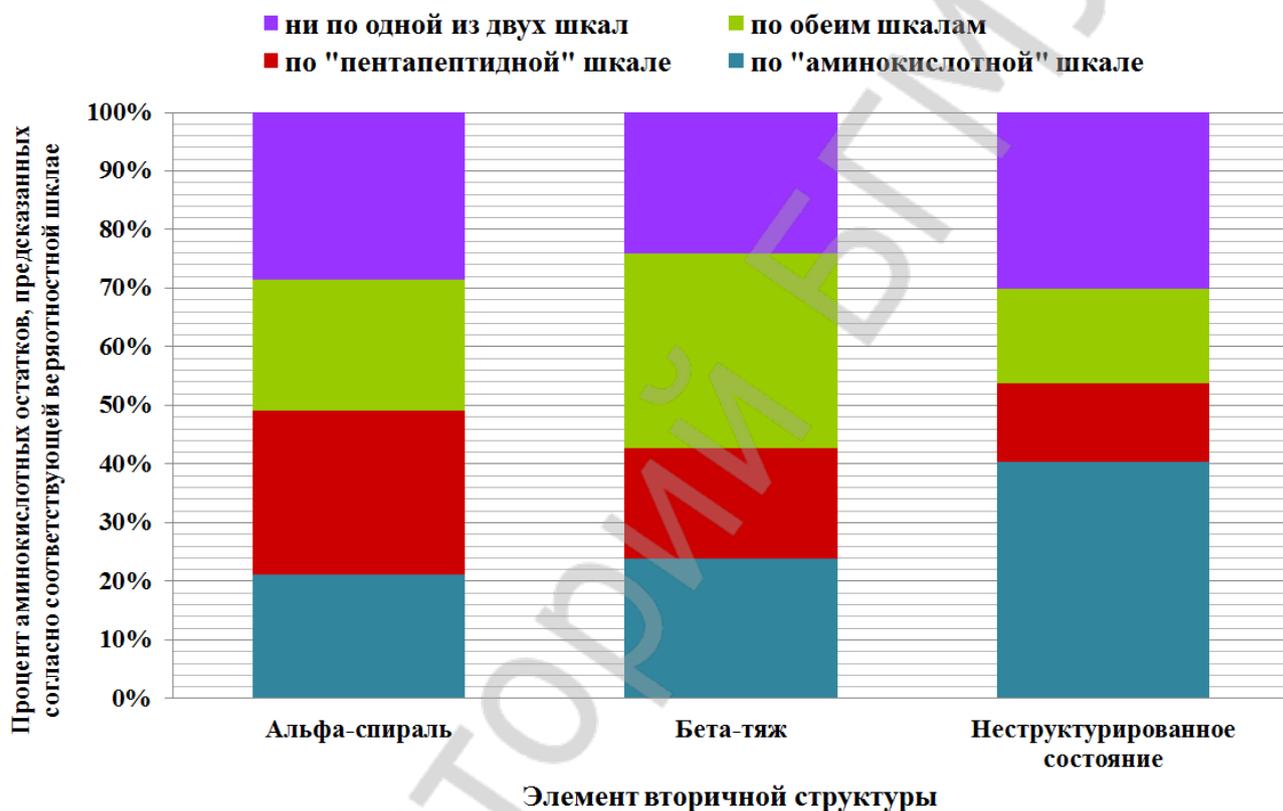


Рис. 1 – Процентное соотношение аминокислотных остатков в альфа-спиралях, бета-тяжах и неструктурированных участках, предсказанных по «аминокислотной» и «пентапептидной» вероятностным шкалам

Для неструктурированных участков наиболее важным фактором формирования является природа аминокислотных остатков, а отнюдь не их чередование. Процент остатков из «соединительных мостиков» между бета-тяжами и альфа-спиралями, находящихся в кластерах из «разрушителей» вторичной структуры, не находящихся в характерных комбинациях, составил 40,28% (Рисунок 1). Действительно, преимущественно гидрофильные пентапептиды начинают играть существенную роль в формировании неструктурированных участков только в белках, кодируемых очень GC-бедными генами [3]. В нашей выборке процент альфа-спиральных или бета-структурных аминокислот, находящихся в комбинациях, характерных для неструктурированного состояния, составил всего 13,57%. Невозможно объяснить неучастие в формировании вторичной структуры для 30,06% аминокислот.

Полученные результаты способствуют совершенствованию критериев отбора фрагментов белков вирусных и бактериальных возбудителей инфекционных болезней для включения их в состав коротких вакцинных пептидов. Для воспроизведения (с высокой долей вероятности) бета-тяжа в составе короткого пептида необходимо совпадение результатов по обеим шкалам алгоритма PentaFOLD. Для воспроизведения альфа-спирали достаточно положительного результата хотя бы по одной из двух вероятностных шкал. О существовании данного фрагмента пептида в неструктурированном состоянии гораздо более достоверно свидетельствуют показания «аминокислотной», а не «пентапептидной» шкалы алгоритма PentaFOLD.

Выводы: 1. Объяснить нахождение в том или ином элементе вторичной структуры с помощью хотя бы одной из двух вероятностных шкал можно лишь для 71,64% аминокислотных остатков: в остальных случаях вторичная структура формируется, в основном, под влиянием взаимодействий между фрагментами белка, отдалёнными друг от друга в первичной аминокислотной последовательности. 2. Для формирования альфа-спиралей более важную роль играют специфические чередования аминокислотных остатков с гидрофобными и гидрофильными боковыми цепями, а не другие структурные особенности их радикалов. 3. Для образования бета-тяжей предпочтительно сочетание двух факторов: специфических чередований гидрофобных и гидрофильных остатков и характерной природы их боковых цепей. 4. Неструктурированные фрагменты белков, как правило, возникают в местах скопления специфических аминокислотных остатков, роль же чередований гидрофильных и гидрофобных аминокислот при этом не так велика.

Литература

1. Adrover, M. Prion fibrillization is mediated by a native structural element that comprises helices H2 and H3 / M. Adrover, K. Pauwels, S. Prigent, C. de Chiara, Z. Xu, C. Chapuis, A. Pastore, H. Rezaei // *J. Biol. Chem.* – 2010. – Vol. 285. – P. 21004-21012.
2. Chou, P. Y. Prediction of the secondary structure of proteins from their amino acid sequence / P. Y. Chou, G. D. Fasman // *Adv. Enzymol. Relat. Areas. Mol. Biol.* – 1978. – Vol. 47. – P. 45-48.
3. Khrustalev, V. V. Stabilization of secondary structure elements by specific combinations of hydrophilic and hydrophobic amino acid residues is more important for proteins encoded by GC-poor genes / V. V. Khrustalev, E. V. Barkovsky // *Biochimie.* – 2012. – Vol.94. – Issue 12. – P. 2706-2715.
4. Khrustalev, V. V. The influence of flanking secondary structures on amino acid content and typical lengths of 3/10 helices / V. V. Khrustalev, E. V. Barkovsky, T. A. Khrustaleva // *Int. J. Proteomics.* – 2014. – Vol. 2014. – P. 1-13.
5. Lim, V. I. Stereochemical theory of the secondary structure of globular proteins. II. Methods of localization of alpha-helical and beta-structural regions / V. I. Lim // *Biofizica.* – 1974. – Vol. 19. – P. 562-575.