

## ВЗАИМОСВЯЗЬ МЕЖДУ ЗНАКОМ УГЛА ОМЕГА ОСТОВА ПОЛИПЕПТИДНОЙ ЦЕПИ И ТИПОМ БОКОВЫХ РАДИКАЛОВ АМИНОКИСЛОТНЫХ ОСТАТКОВ

© 2017 г. И.Ю. Торшин, А.В. Батяновский\*, Л.А. Урошлев\*\*,  
Н.Г. Есипова\*\*, В.Г. Туманян\*\*

*Химический факультет Московского государственного университета имени М.В. Ломоносова,  
119991, Москва, Ленинские горы, 1/3*

*\*Институт биофизики и клеточной инженерии НАН Беларуси, 220072, Минск, ул. Академическая, 27, Беларусь*

*\*\*Институт молекулярной биологии им. В.А. Энгельгардта РАН, 119991, Москва, ул. Вавилова, 32*

*E-mail: tuman@imb.ru*

Поступила в редакцию 02.12.16 г.

Показано, что угол  $\omega$ , характеризующий неплоскостность пептидной группы, является параметром, существенным для описания конформаций остова полипептидной цепи в контексте конкретных боковых радикалов аминокислот. Обнаружены кластеры конформаций, соответствующие ранее найденным конформационно-стабильным пептидам.

*Ключевые слова: полипептидная цепь, конформации, карта Рамачандрана, кластеризация, боковые радикалы аминокислот, конформационно-стабильные пептиды.*

Карта Рамачандрана [1] – удобный, широко используемый инструмент для характеристики конформаций основной цепи белков. На картах Рамачандрана отражаются значения двугранных углов, задающих повороты вокруг связей, примыкающих к  $C^\alpha$ -атомам (углы  $\phi$ ,  $\psi$ ; рис. 1). В то же время, кроме углов  $\phi$  и  $\psi$ , в основной цепи существует еще один двугранный угол (угол  $\omega$ , рис. 1). Этот угол, отклонения которого от значений  $180^\circ$  и  $0^\circ$  придают пептидной группе неплоский характер, достаточно редко принимается во внимание при анализе конформаций. Логично расширить размерность карты Рамачандрана путем добавления недостающего двугрannого угла  $\omega$ .

Важность угла  $\omega$  для описания конформации полипептидной цепи следует из теории химической связи. Как известно,  $sp^2$ -гибридизация атомов N и  $C'$ , образующих пептидную связь, в сильной степени тормозит вращение вокруг оси  $C'-N$ , что позволяет говорить о принципиальной «плоскостности» пептидной группы. Однако постулирование плоскостности пептидной группы является всего лишь приближенным описанием ее конформации. Для изучения отклонений пептидной группы от плоскости привлекаются теоретические и экспериментальные методы [2–8]. В целом экспериментальные данные и теоретические соображения указывают

на возможность реализации неплоских деформаций пептидной группы [2,3], которые достаточно хорошо описываются изменением угла  $\omega$ . Для еще более полного описания конформации необходимо также учитывать пирамидализацию атомов  $C'$  и N [4–6]. Учет значений угла  $\omega$  необходим, в частности, для объяснения стабилизации определенных типов вторичной структуры (например,  $\alpha$ -спирали и левой спирали типа РРІІ за счет  $n \rightarrow \pi^*$ -взаимодействия между  $C'=O$ -группой и атомом  $C'$  следующей по цепи  $C'=O$ -группы [9–11]).

В связи с увеличением числа белковых структур с весьма высоким разрешением (по крайней мере не хуже  $2,0 \text{ \AA}$ ) можно утверждать, что отклонения значения угла вращения  $\omega$  вокруг пептидной связи устанавливаются с точностью, достаточной для того, чтобы можно было строить карту Рамачандрана, включая угол  $\omega$  в качестве одной из переменных. В силу того, что наблюдаются лишь незначительные (около  $5^\circ$ ) отклонения угла  $\omega$  от идеального значения ( $180^\circ$ ), достаточно описывать эти отклонения только знаками плюс или минус, что качественно определяет характер деформации пептидной плоскости.

В настоящем исследовании был проведен анализ взаимосвязей между знаком угла  $\omega$  основной цепи и типом боковых цепей аминокислот.

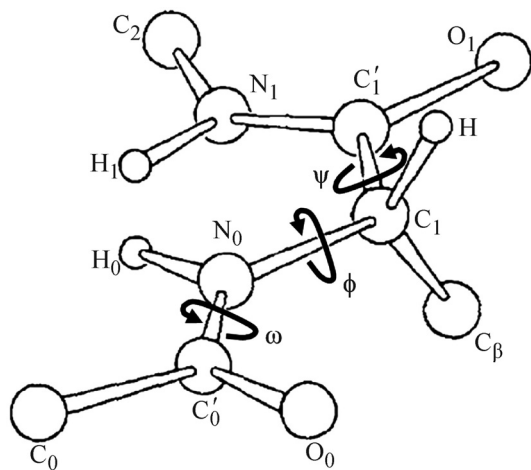


Рис. 1. Двугранные углы в полипептидной цепи. Показана «дипептидная единица» Рамачандрана [1].

кислотных остатков. Анализ проводили в четыре этапа следующим образом.

## МАТЕРИАЛЫ И МЕТОДЫ

На первом этапе на основании данных PDB-банка (релиз 2016 г.) была сформирована выборка из структур 121450 белковых цепей, взятых из 62096 PDB-файлов (исключены цепи с идентичными последовательностями и с разрешением, худшим чем 2,0 Å). Значения двугранных углов  $\phi$ ,  $\psi$  и  $\omega$  были рассчитаны для 52563104 аминокислотных остатков с известными координатами каждого неводородного атома основной цепи.

На втором этапе для каждого типа аминокислоты была построена соответствующая карта Рамачандрана в координатах  $\phi$  и  $\psi$ . Карты строили как без учета знака угла  $\omega$ , так и по отдельности для положительных и отрицательных значений этого угла. В результате было получено 60 карт (20 – без учета знака угла  $\omega$ , 20 – для  $\omega \geq 0$  и 20 – для  $\omega < 0$ ).

На третьем этапе были оценены «расстояния» между полученными картами Рамачандрана для различных аминокислот. В качестве функции расстояния между двумерными эмпирическими функциями распределения значений  $\phi$  и  $\psi$ , соответствующими картам Рамачандрана для индивидуальных аминокислот, была использована информационная метрика Фишера.

На четвертом этапе после вычисления всех попарных расстояний между картами была проведена процедура кластеризации. Для кластеризации конформаций использовали высокочувствительный алгоритм поиска метрических

сгущений точек (кластеров близко лежащих точек с высокой плотностью точек [12]).

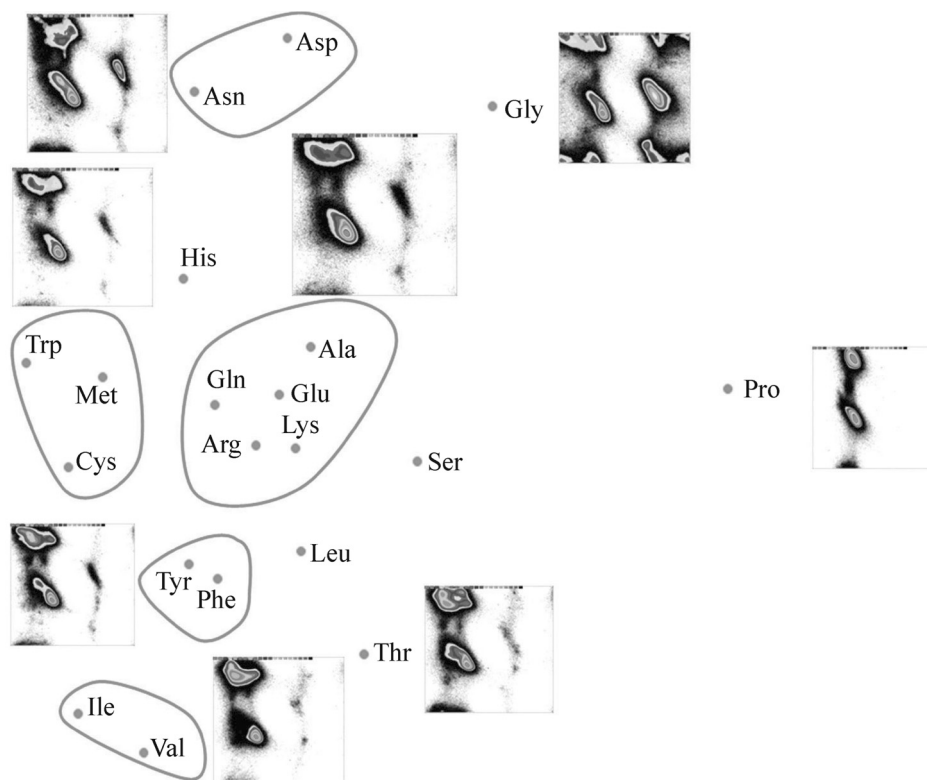
## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В результате кластеризации карт Рамачандрана для индивидуальных аминокислот (без учета знака угла  $\omega$ ) мы установили схожесть карт для следующих кластеров аминокислот: {Asn, Asp}, {Tyr, Phe}, {Met, Cys, Trp}, {Ile, Val} и {Ala, Glu, Gln, Arg, Lys}. Остальные аминокислоты не входили в данные кластеры и образовывали одноточечные кластеры {Gly}, {His}, {Pro}, {Ser}, {Thr}, {Leu}, т.е. характеризовались особыми распределениями точек на картах Рамачандрана, отличными от распределений для всех остальных аминокислот (рис. 2). В определенном смысле полученные кластеры можно рассматривать как некоторые «классы конформационных состояний» главных цепей аминокислот.

Естественно искать сходство и различие полученной в настоящей работе классификации аминокислот с многочисленными вариантами такой классификации, имеющейся в литературе. Результаты использования двух методов кластеризации (анализа главных компонент и иерархической кластеризации) приводят авторов работы [13] к сходным выводам. Так, из анализа главных компонент матриц сходства, установленных на основании статистического анализа аминокислотных последовательностей, следует, что аминокислоты Val, Ile, Leu и Met и Phe кластеризуются с Phe, Tyr и Trp; а остатки Ser, Thr и Ala плюс Asp, Asn и Glu, Lys и Arg образуют ядро второго большого кластера. Такие аминокислоты, как Gly, His и Pro, занимают обособленное положение. Авторы работы [13] (см. также работу [14]), исходя из иерархической кластеризации, отмечают различие между большими аминокислотными остатками, объединенными с неполярными или умеренно полярными остатками (Val, Ile, Leu, Met, Phe, Tyr и Trp) и остатками, которые являются малыми неполярными, полярными и заряженными. Особое положение также занимают His, Pro и Gly, к ним присоединяется и Trp.

Таким образом, не усматривается сходства полученной в работе [13] картины кластеризации с результатами, полученными с помощью нашего метода, за исключением «одиночных» кластеров, включающих Gly, His и Pro. Дополнительно отметим, что кластер Tyr-Phe наблюдается как нами, так и в ряде других работ.

Представляет интерес сравнить наши результаты с данными, полученными в работе [15] методом оценки частот разных пар остат-



**Рис. 2.** Результаты кластеризации карт Рамачандрана для отдельных типов аминокислотных остатков (без ограничений на значения угла  $\omega$ ), с типичными примерами.

ков, разделенных разным расстоянием вдоль полипептидной цепи белка. Кластер Asp, Asn из работы [15] совпадает с аналогичным кластером на рис. 2. Отметим ограниченное сходство кластера Val, Ile, Leu из работы [15] и кластера Val, Ile, полученного в настоящей работе. Далее, кластер Lys, Arg, Gln + Glu, полученный в работе [15], напоминает кластер Ala, Lys, Arg, Gln, Glu (см. рис. 2). Однако принципиальный момент состоит в отсутствии Ala в первом случае. Наконец, одноточечные кластеры в работе [15] представлены остатками Gly, His, Pro, Trp, Cys, Met; в настоящей работе – остатками, согласно нашим данным, Gly, His, Pro, Ser, Leu, Thr.

Наиболее существенный результат состоит в том, что кластеризация карт Рамачандрана, полученных с учетом знака угла  $\omega$ , указывает на существенные отличия в конформации остова полипептидной цепи для одной и той же аминокислоты при противоположных знаках угла  $\omega$ . Так, при учете знака угла  $\omega$  кластер {Ala, Glu, Gln, Arg, Lys} на рис. 2 распался на два кластера: {Ala ( $\omega \geq 0$ ), Glu ( $\omega \geq 0$ ), Gln ( $\omega \geq 0$ ), Arg ( $\omega \geq 0$ ), Lys ( $\omega \geq 0$ )} и {Glu ( $\omega < 0$ ), Gln ( $\omega < 0$ ), Arg ( $\omega < 0$ ), Lys ( $\omega < 0$ )} на рис. 3, причем аланин вышел из состава кла-

стера при  $\omega < 0$ . Кластер {Asn, Asp} распался на {Asn ( $\omega \geq 0$ ), Asp ( $\omega \geq 0$ )} и {Asn ( $\omega < 0$ ), Asp ( $\omega < 0$ )}, кластер {Tyr, Phe} – на {Tyr ( $\omega \geq 0$ ), Phe ( $\omega \geq 0$ )} и {Tyr ( $\omega < 0$ ), Phe ( $\omega < 0$ )}. За исключением пролина, образовавшего двухточечный кластер {Pro ( $\omega \geq 0$ ), Pro ( $\omega < 0$ )}, все остальные аминокислоты образовывали одноточечные кластеры {Gly ( $\omega \geq 0$ )}, {Gly ( $\omega < 0$ )}, {His ( $\omega \geq 0$ )}, {His ( $\omega < 0$ )}, {Trp ( $\omega \geq 0$ )}, {Trp ( $\omega < 0$ )}, {Met ( $\omega \geq 0$ )}, {Met ( $\omega < 0$ )}, {Cys ( $\omega \geq 0$ )}, {Cys ( $\omega < 0$ )}, {Val ( $\omega \geq 0$ )}, {Val ( $\omega < 0$ )}, {Ile ( $\omega \geq 0$ )}, {Ile ( $\omega < 0$ )}, {Ser ( $\omega \geq 0$ )}, {Ser ( $\omega < 0$ )}, {Thr ( $\omega \geq 0$ )}, {Thr ( $\omega < 0$ )}, {Leu ( $\omega \geq 0$ )}, {Leu ( $\omega < 0$ )} (рис. 3).

Таким образом, можно утверждать, что за исключением пролина отражаемые на карте Рамачандрана наборы возможных конформаций остова полипептидной цепи существенно отличаются для одной и той же аминокислоты с различными знаками угла  $\omega$ . Например, карта Рамачандрана для Ala ( $\omega \geq 0$ ) гораздо ближе к картам для Glu ( $\omega \geq 0$ ), Gln ( $\omega \geq 0$ ) и др. (эти карты принадлежат одному кластеру {Glu ( $\omega \geq 0$ ), Gln ( $\omega \geq 0$ ), Arg ( $\omega \geq 0$ ), Lys ( $\omega \geq 0$ )}, чем к карте для Ala ( $\omega < 0$ ).

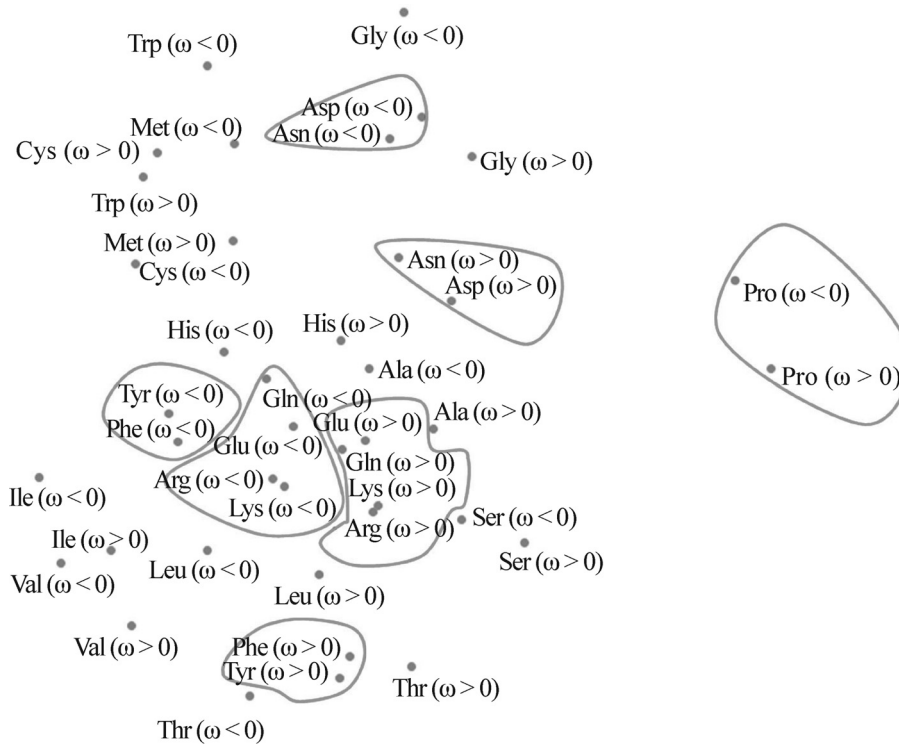


Рис. 3. Результаты кластеризации карт Рамачандрана для отдельных типов аминокислотных остатков с учетом значений угла  $\omega$ .

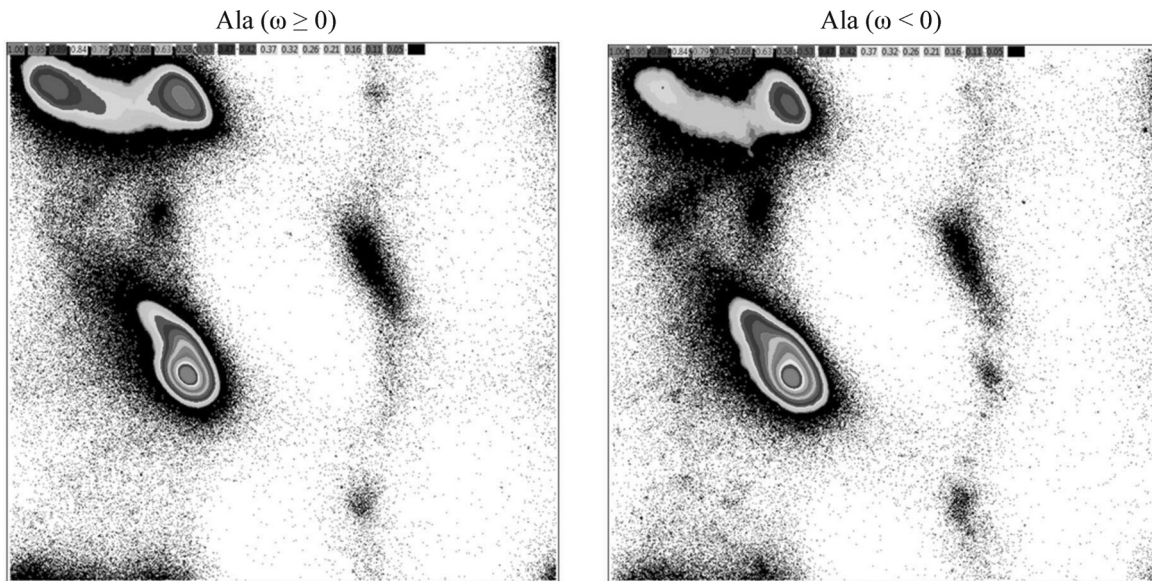


Рис. 4. Карты Рамачандрана с учетом плотности точек и знака угла  $\omega$  для аланина.

Примеры отличий между картами Рамачандрана для индивидуальных аминокислот с различными знаками  $\omega$  приведены на рис. 4-7.

Карты Рамачандрана (без учета угла  $\omega$ ) для индивидуальных аминокислот на основе экспе-

риментальных данных были получены в работе [16]. Наблюдается качественное сходство карт, построенных нами, с этими картами. Проведение более строгого сравнения затруднительно, так как авторами [16] не было проведено цветовое шкалирование полученных данных и, тем

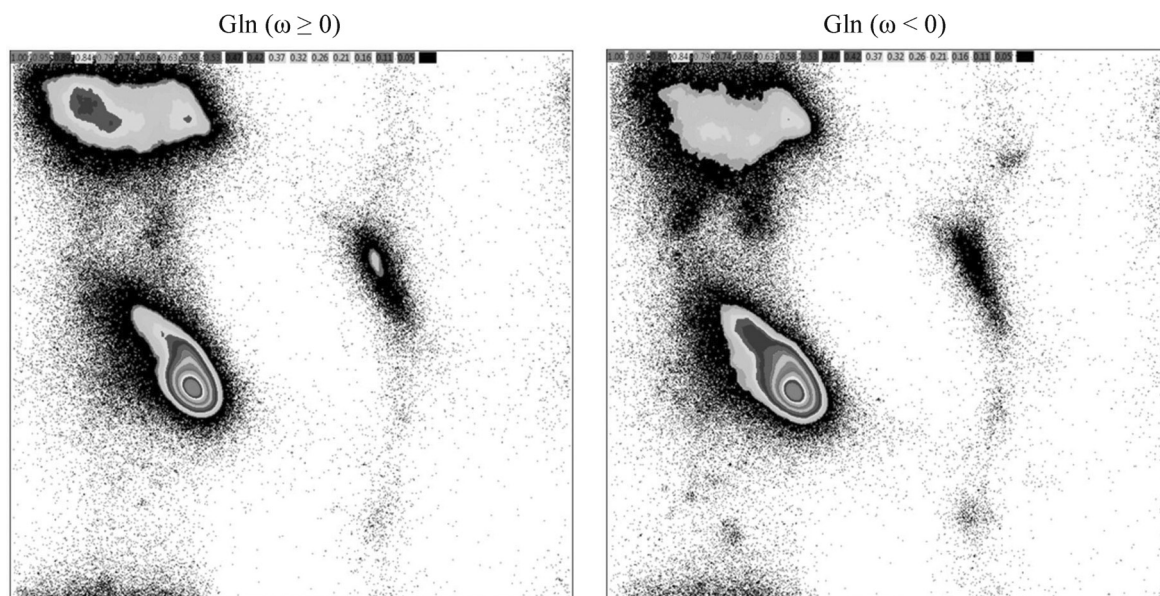


Рис. 5. Карты Рамачандрана с учетом плотности точек и знака угла  $\omega$  для глутамина.

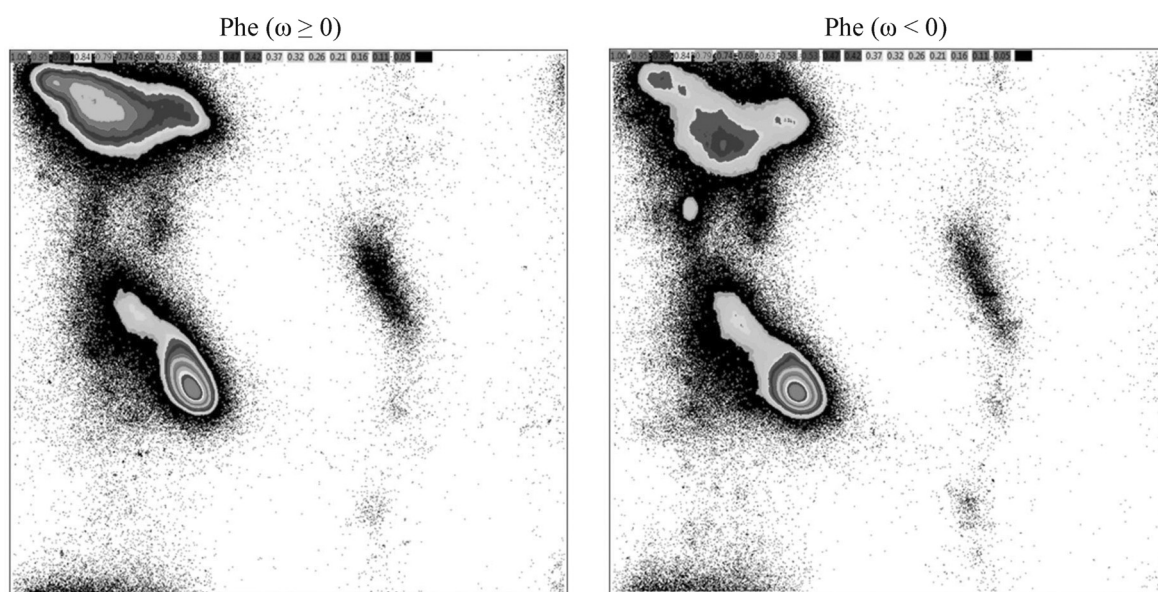


Рис. 6. Карты Рамачандрана с учетом плотности точек и знака угла  $\omega$  для фенилаланина.

самым, представленная в этой работе информация оказалась неполной.

В связи с проведенным рассмотрением взаимосвязи между типом аминокислотного остатка и конформацией основной цепи (картами углов  $\phi$  и  $\psi$  соответствующей дипептидной единицы с учетом или без учета угла  $\omega$ ) представляет интерес привлечь для дальнейшего обсуждения данные по исследованным нами конформационно-стабильным сегментам полипеп-

тидной цепи. В работах [17–19] была получена выборка тетрапептидов, конформация которых в большей степени определяется именно аминокислотной последовательностью и в меньшей степени – контекстом тетрапептида в тех или иных белках. Степень близости конформаций таких пептидов достаточно высока: отличие по двугранным углам не превосходит  $10^\circ$ . Список из 943 таких «конформационно-стабильных» тетрапептидов приведен в работе [19]. В отличие от нашего подхода в исследованиях, развивае-

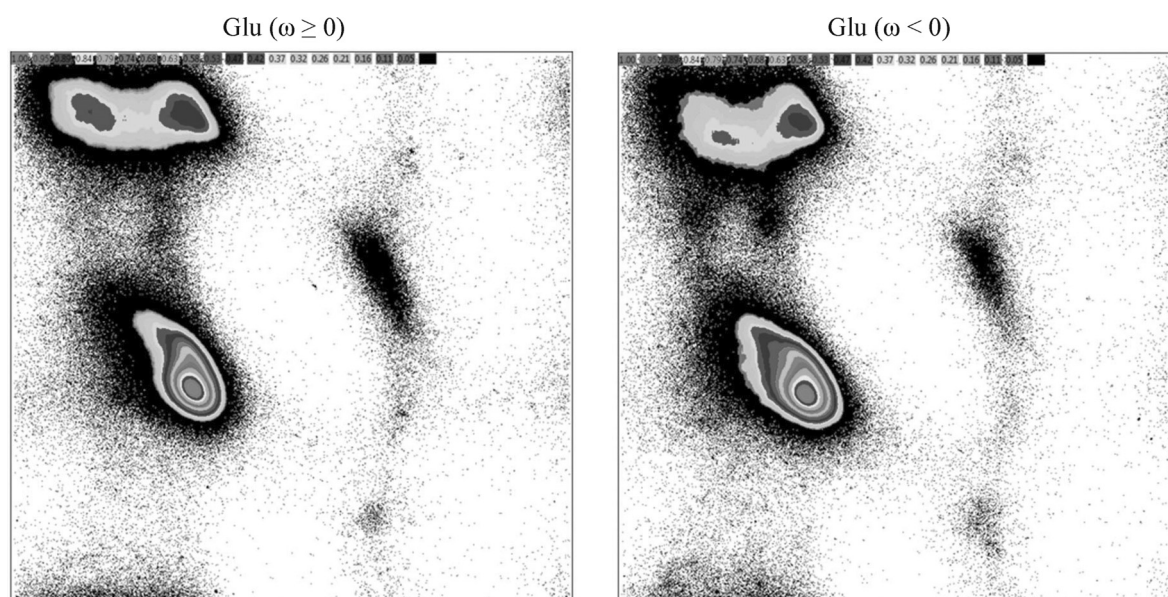


Рис. 7. Карты Рамачандрана с учетом плотности точек и знака угла  $\omega$  для глутаминовой кислоты.

мых Д. Бэйкером с соавторами [20], требования к близости конформационных состояний олигопептидов с одинаковой последовательностью являются существенно более мягкими.

В таблице приведены отличия в аминокислотном составе конформационно-стабильных тетрапептидов и, для сравнения, для всех 160000 тетрапептидов; для вычисления состава исполь-

Аминокислотный состав конформационно-стабильных тетрапептидов и всех тетрапептидов (на основе выборки, приведенной в работе [19])

Аминокислота	Конформационно-стабильные тетрапептиды	Все тетрапептиды
<b>Ala</b>	<b>0,246</b>	<b>0,126</b>
<b>Cys</b>	<b>0,002</b>	<b>0,010</b>
Glu	0,137	0,089
<b>Asp</b>	<b>0,028</b>	<b>0,051</b>
<b>Gly</b>	<b>0,003</b>	<b>0,034</b>
<b>Phe</b>	<b>0,022</b>	<b>0,039</b>
Ile	0,064	0,059
<b>His</b>	<b>0,005</b>	<b>0,020</b>
Lys	0,071	0,065
<b>Met</b>	<b>0,013</b>	<b>0,022</b>
Leu	0,164	0,117
<b>Asn</b>	<b>0,010</b>	<b>0,032</b>
Gln	0,061	0,049
<b>Pro</b>	<b>0,001</b>	<b>0,021</b>
<b>Ser</b>	<b>0,011</b>	<b>0,046</b>
Arg	0,076	0,060
<b>Thr</b>	<b>0,018</b>	<b>0,043</b>
<b>Trp</b>	<b>0,004</b>	<b>0,015</b>
Val	0,047	0,064
<b>Tyr</b>	<b>0,014</b>	<b>0,035</b>

Примечание. Выделены статистически достоверные отличия в частотах встречаемости (критерий  $\chi^2$ ).

зовали выборку олигопептидов из работы [18]. Можно видеть, что в аминокислотном составе конформационно-стабильных тетрапептидов в наибольшей степени представлены аминокислоты, характерные для  $\alpha$ -спирали: доля этих аминокислот существенно повышена (Ala – в два раза по сравнению с полной выборкой, Leu и Glu и др.). В то же время отмечено существенное снижение содержания «конформационно-дестабилизирующих» аминокислот пролина (в 20 раз), глицина (в 10 раз) и гистидина (в 4 раза).

Таким образом, аминокислотный состав конформационно-стабильных пептидов обогащен «стабилизирующими» аминокислотами Ala, Leu, Glu, Arg, Lys, Ile, Gln. Почти все эти аминокислоты (за исключением Leu) соответствуют основному кластеру Ala, Glu, Arg, Lys, Gln, аккумулирующему «альфа-спиральные» аминокислоты (см. рис. 2). Этот кластер расщепляется на два при учете знака угла  $\omega$  (рис. 3), причем при  $\omega > 0$  состав кластера не меняется, а при  $\omega < 0$  от этого кластера отщепляется аланин. Как известно, аланин – аминокислота, имеющая наибольшее сродство к  $\alpha$ -спирали, поэтому кластер {Glu ( $\omega < 0$ ), Gln ( $\omega < 0$ ), Arg ( $\omega < 0$ ), Lys ( $\omega < 0$ )} на рис. 3 уже нельзя назвать вполне «альфа-спиральным».

Таким образом, анализ конформаций основной цепи, выполненный с учетом знака угла  $\omega$ , отражающего направления «скручивания» полипептидной плоскости, позволил установить кластеры конформаций отдельных аминокислот. Показано соответствие между этими кластерами конформаций и особенностями аминокислотного состава выделенных нами ранее «конформационно-стабильных пептидов».

Полученные данные о конформационных свойствах остова полипептидной цепи в контексте конкретных аминокислот с учетом вариации угла  $\omega$  представляют общий интерес и должны приниматься во внимание при анализе специфических ДНК-белковых взаимодействий. В частности, кластеризация аминокислот по конформационным свойствам, в том числе таким, как конформационная стабильность, естественным образом соотносится со сродством аминокислот к основаниям ДНК.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проекты №№ 14-04-00639, 16-54-00219-Бел и 15-04-99605) и Белорусского республиканского фонда фундаментальных исследований (проект № Б16Р-185).

## СПИСОК ЛИТЕРАТУРЫ

1. G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, *J. Mol. Biol.* **7**, 95 (1963).
2. L. Bednarova, P. Malon, and P. Bour, *Chirality* **19**, 775 (2007).
3. S. Yamada, *J. Org. Chem.* **61**, 941 (1996).
4. G. A. Jeffrey, K. N. Houk, M. N. Paddon-Row, et al., *J. Am. Chem. Soc.* **107**, 321 (1985).
5. J. D. Dunitz and F. K. Winkler, *Acta Crystallogr. B* **31**, 251 (1975).
6. L. Esposito, L. Vitagliano, F. Zagari, and L. Mazzarella, *Protein Sci.* **9**, 2038 (2000).
7. G. N. Ramachandran and A. S. Kolaskar, *Biochim. Biophys. Acta* **303**, 385 (1973).
8. M. W. MacArthur and J. M. Thornton, *J. Mol. Biol.* **264**, 1180 (1996).
9. R. W. Newberry, S. J. Orke, and R.T. Raines, *Org. Lett.* **18**, 3614 (2016).
10. Ch. Fufesan, *Proteins* **78**, 2831 (2010).
11. M. D. Shoulders and R. T. Raines, *Annu. Rev. Biochem.* **78**, 929 (2009).
12. I. Y. Torshin, *Pattern Recognition and Image Analysis* **26**, 483 (2016).
13. M. S. Johnson and J. P. Overington, *J. Mol. Biol.* **233**, 716 (1993).
14. A. Prlie, F. S. Domingues, and V. J. Sippl, *Protein eng.* **13**, 545 (2000).
15. S. I. Rogov and A. N. Nekrasov, *Protein eng.* **14**, 459 (2001).
16. S. A. Hollingsworth and P. A. Karplus, *Biomol. Concepts* **1**, 271 (2010).
17. А. В. Батяновский и П. К. Власов, *Биофизика* **53**, 556 (2008).
18. А. В. Батяновский, Н. Г. Есипова и С. Э. Шноль, *Биофизика* **54**, 1137 (2009).
19. А. В. Батяновский, И. Д. Волотовский, В. А. Намиот и др., *Биофизика* **60**, 437 (2015).
20. C. Bystroff, K. T. Simons, K. F. Han, and D. Baker, *Curr. Opin. Biotechnol.* **7**, 417 (1996).

## **Interrelation Between the Sign of Polypeptide Backbone Omega Angle and the Type of Side Chain Radicals of Amino Acid Residues**

**I.Yu. Torshin\***, **A.V. Batyanovskii\*\***, **L.A. Uroshlev\*\*\***,  
**N.G. Esipova\*\*\***, and **V.G. Tumanyan\*\*\***

*\*Department of Chemistry, Lomonosov Moscow State University, Leninskie Gory 1/3, Moscow, 119991 Russia*

*\*\*Institute of Biophysics and Cell Engineering, National Academy of Sciences of Belarus,  
ul. Akademicheskaya 27, Minsk, 220072 Belarus*

*\*\*\*Engelhard Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, Moscow, 119991 Russia*

It was shown that dihedral angle  $\omega$  manifesting nonplanarity of the peptide group is a parameter which is essential for a description of conformations of the polypeptide chain backbone in the context of definite side chain radicals of amino acid residues. The conformational clusters which correspond to previously shown conformationally stable peptides are observed.

*Key words: polypeptide chain, conformations, Ramachandran plot, clustering, side chain radicals of amino acid residues, conformationally stable peptides*