

УДК 577.3

ИССЛЕДОВАНИЕ β -ИЗГИБОВ В ГЛОБУЛЯРНЫХ БЕЛКАХ

© 2005 г. С.Р. Амирова, Ю.В. Мильчевский, И.В. Филатов, Н.Г. Есипова, В.Г. Туманян

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

Поступило в редакцию 01.09.05 г.

В развиваемом нами новом подходе к предсказанию β -изгибов на основе статистического метода дискриминантного анализа и конформационных расчетов модельных структур не используется множественное выравнивание последовательностей для поиска гомологов, которые, кроме всего, не всегда можно найти. Он основывается на анализе полипептидной цепи того белка, для которого делается предсказание. Важное преимущество нашего подхода заключается в возможности выявления свойств белка (локальных и глобальных), влияющих на образование β -изгибов. С целью изучения энергетических факторов, способствующих образованию β -изгибов в белках, мы использовали компоненты конформационной энергии, а также различные функции этих величин.

Ключевые слова: вторичные структуры белков, β -изгибы – предсказания, конформационная энергия, дискриминантный анализ.

Изгибы полипептидной цепи в глобулярных белках – важный элемент трехмерной структуры, обеспечивающий специфичность топологии белковых фолдов [1]. Один из наиболее часто встречающихся типов изгиба – это β -изгиб.

Элемент супервторичной структуры β -изгиб представляет собой четыре последовательных аминокислотных остатков, при этом расстояние между первым и четвертым по цепи остатками должно быть меньше, чем 7 Å. Различные β -изгибы можно классифицировать в рамках девяти типов, в зависимости от значений двугранных углов ϕ и ψ второго и третьего аминокислотных остатков в β -изгибе [2]. Существуют различные подходы к предсказанию β -изгибов в белках [3–6].

Моделирование β -изгибов и расчет компонентов конформационной энергии проводили методами молекулярной механики с использованием компьютерной программы ICM [7]. Для расчетов были отобраны 190 белковых цепей глобулярных белков из Protein Data Bank, гомология которых была менее 35% [2,3].

Процедура получения компонентов конформационной энергии заключалась в следующем: аминокислотная последовательность рассматриваемых белковых цепей была разбита на сегменты. Каждый следующий сегмент был полу-

чен путем сдвига по цепи на одну аминокислоту. Размер сегмента составлял 14 аминокислотных остатков. В данной работе приводятся результаты конформационного анализа для β -изгиба I-го типа как наиболее распространенного. Для стандартизации окружения β -изгиба пяти остаткам слева и пяти справа от центральных была задана конформация антипараллельной β -структуры. Для сближения шестого и девятого остатков на расстояние, меньшее чем 7 Å, вводили стягивающие потенциалы. Для нахождения оптимальной конформации боковых радикалов для всех аминокислотных остатков анализируемого сегмента применялся метод редукции библиотеки ротамеров [8].

Методом молекулярной механики с использованием метода Монте-Карло была найдена оптимальная конформация для анализируемого сегмента и получены характеризующие это состояние восемь компонентов конформационной энергии: полная конформационная энергия, ван-дер-ваальсова энергия, энергия водородных связей, электростатическая энергия, торсионная энергия, энергия деформации валентных связей, энергия деформации фазовых углов, энергия стягивающих потенциалов. Подобная процедура расчета компонентов энергии была выполнена для каждого из сегментов всех рассматриваемых белковых цепей.

Полученные компоненты конформационной энергии, а также их квадраты и кубы, рассматривались в качестве входных переменных для программы дискриминантного анализа. На основании этих переменных и последовательности аминокислот белка производилось предсказание β -изгибов в пространственной структуре. Для выявления β -изгибов в белках с известной пространственной структурой использовалась система классификации STRIDE [9]. Использование тестовой выборки, содержащей 190 белковых цепей с известной пространственной структурой, дало возможность оценить точность предсказания путем сравнения полученных результатов с реальной структурой. В табл. 1 показаны результаты предсказания двух состояний (изгиб–не изгиб) для каждого аминокислотного остатка 190 рассматриваемых белковых цепей

В табл. 1 приведены результаты предсказания β -изгибов I типа. Табл. 1 содержит матрицу точности 2×2 для двух состояний β -изгиб – не β -изгиб: 64% это положительное правильное предсказание, 46% – отрицательное правильное предсказание, 35% – недопредсказание и 55% – перепредсказание. По табл. 1 суммарный критерий точности для двух состояний составил $Q_2 = 51\%$.

Таблица 1. Результаты предсказания β -изгиб – не β -изгиб (%)

Тип вторичной структуры	Любая структура, кроме β -изгиба I типа	β -изгиб I типа
Любая структура, кроме β -изгиба I типа	45	55
β -изгиб I типа	35	65

Для оценки точности предсказания использовалось несколько тестовых белковых выборок, при этом обучающие выборки были исключены. Т.е. во избежание перекрытия обучающих и тестовых выборок был применен тест jack knife [10]. Обучающие и тестовые выборки формировали из 190 рассматриваемых белковых цепей. Вклад компоненты энергии в образование β -изгиба I типа рассчитывался на основании статистической значимости в статистике Фишера для соответствующей входной

переменной в дискриминантном анализе. Компоненты конформационной энергии, внесшие, по данным дискриминантного анализа, наибольший вклад в предсказание β -изгиба I типа, приведены в табл. 2.

Таблица 2. Компоненты энергии, вносящие основной вклад в предсказание β -изгиба

Название компоненты энергии	Статистическая значимость соответствующей переменной
Торсионная энергия	251,134
Энергия водородных связей	95,716
Энергия Ван-дер-Ваальса	21,479

Частный случай для α -спиральных белков рассмотрен в [11].

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты № 05-04-4962 и № 03-04-4917), программы Президиума РАН «Молекулярная и клеточная биология» и контракта с Министерством образования и науки (№ 02.434.11.1008).

СПИСОК ЛИТЕРАТУРЫ

1. Takano K., Yamagata Y., Yutani K. // *Biochemistry*. 2000. V. 39. P. 8655–8665.
2. Hutchinson E.G., Thornton J.M. // *Protein Sci.* 1994. V. 3. P. 2207–2216.
3. Chou K.C., Blinn J.R. // *J. Protein Chem.* 1997. V.16. P. 575–595.
4. Shepherd A-J., Gorse D., Thornton J.M. // *Protein Sci.* 1999. V. 8. P. 1045–1055.
5. Lin T.H., Wang G.M., Wang Y.T. // *J. Chem. Inf. Comput. Sci.* 2002. V. 42. P. 123–133.
6. Harpreet Kaur, G. P. S. Raghava // *Bioinformatics*. 2004. V. 20. P. 2751–2758
7. Abagyan R.A., Totrov M.M., Kuznetsov D.N. // *J. Comp. Chem.* 1994. V.15. P. 488–506.
8. Филатов И.В., Мильчевский Ю.В., Есипова Н.Г., Туманян В.Г. // *Биофизика*. 2004. Т. 49, вып. 6. С. 1144–1146.
9. Frishman D., Argos P. // *Proteins*. 1995. V. 23. P. 566–579.
10. Mardia K.V., Kent J.T., Biddy J.M. // *Academic Press, London*. 1979. P. 322–381.
11. Kilosanidze G.T., Kutsenko A.S., Esipova N.G., Tumanyan V.G. // *FEBS Lett.* 2002. V. 510. P. 13–16.

Study of β -Turns in Globular Proteins

S.R. Amirova, Ju.V. Milchevsky, I.V. Filatov, N.G. Esipova, and V.G. Tumanyan

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, Moscow, 119991 Russia

The formation of β -turns in globular proteins has been studied by the method of molecular mechanics. Statistical method of discriminant analysis was applied to calculate energy components and sequences of oligopeptide segments, and after this prediction of I type β -turns has been drawn. The accuracy of true positive prediction is 65%. Components of conformational energy considerably affecting β -turn formation were delineated. There are torsional energy, energy of hydrogen bonds, and van der Waals energy.

Key words: protein secondary structure, β -turn prediction, conformational energy, discriminant analysis

Сдано в набор 25.10.05	Подписано к печати 19.12.05	Формат бумаги 60x88 $\frac{1}{8}$		
Офсетная печать	Усл. печ. л. 24,0	Усл. кр.-отг. 7,4 тыс.	Уч.-изд. л. 24,1	Бум. л. 12,0
	Тираж 300 экз.	Зак. 908		

Учредители: Российская академия наук,
Институт биофизики клетки РАН

Издатель – Научно-производственное объединение «Издательство «Наука»»
117997, Москва, Профсоюзная ул., 90

Отпечатано в ППП «Типография «Наука», 121099, Москва, Шубинский пер., 6
