



УДК 577.112.083.3 : 519.68

МЕТОД ПОИСКА АМФИПАТИЧНЫХ СТРУКТУР В БЕЛКОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЯХ

Габриэлян А.Э., Иванов В.С., Кожич А.Т.

Институт биоорганической химии им. М. М. Шемякина Академии наук СССР, Москва

В результате интенсивно ведущихся исследований по поиску участков в белковых антигенах, которые узнаются Т-клетками, к настоящему времени выяснена структура более 40 Т-клеточных эпитопов. Но используемые экспериментальные методы весьма трудоемки, поэтому возможность теоретического предсказания иммунодоминантных Т-эпитопов позволила бы значительно ускорить работу по их поиску.

При сравнительном анализе аминокислотных последовательностей ряда известных Т-эпитопов была обнаружена закономерность, присущая большинству из них [1]. В этих пептидах чередуются гидрофильные и гидрофобные аминокислоты, причем таким образом, что в случае спиральной конформации такого пептида данная спираль была бы амфипатичной, т. е. одна ее сторона была бы преимущественно гидрофильной, а противоположная — преимущественно гидрофобной. Для структуры Т-эпитопов характерно также низкое содержание аминокислот, препятствующих образованию α -спирали. Эти факты позволили выдвинуть гипотезу о том, что Т-эпитопы являются амфипатичными спиральями.

Был предложен ряд программ поиска Т-эпитопов как участков белковой последовательности с периодическим чередованием гидрофильных и гидрофобных аминокислот [2—4]. Наиболее популярна программа АМРН1 [2], ориентированная на ЭВМ VAX 11/780, но ее использование на лабораторных персональных компьютерах сопряжено с большими затратами расчетного времени.

Нами разработан новый алгоритм поиска участков белковой последовательности с периодическим чередованием гидрофильных и гидрофобных аминокислот. Программа, использующая этот алгоритм, при значительно меньшем времени счета и требуемой памяти ЭВМ не уступает по предсказательной способности программе АМРН1 и может быть реализована на любом персональном или карманном компьютере.

Для оценки периодичности на выбранном участке белковой последовательности рассчитывается коэффициент парной корреляции [5] между значениями гидрофобностей аминокислот, составляющих этот участок, и значениями периодической функции $y = \cos x$. Коэффициент корреляции рассчитывался по формуле

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}, \quad (1)$$

где x_i — гидрофобность i -й аминокислоты на выбранном участке; y_i — соответствующее значение периодической функции, которое вычисляется по формуле

$$y_i = \cos((i - 1) \cdot P), \quad (2)$$

где i — порядковый номер аминокислоты на выбранном участке ($i = 1, 2, 3, \dots$); P — период (120° для 3_{10} -спирали, 90° для α -спирали, 180° для β -складчатого слоя; \bar{x} , \bar{y} — средние значения гидрофобности и функции $\cos x$ на выбранном участке.

При периодическом чередовании гидрофильных и гидрофобных аминокислот на рассматриваемом участке коэффициент корреляции с периодической функцией будет выше, чем при их случайном распределении. Расчеты проводились для отрезка в 11 аминокислот, затем рамка сдвигалась на 1 аминокислоту. Гидрофобности аминокислотных остатков оценивались по шкале Паркера [6]. Если рассматриваемый отрезок содержал подряд 4 гидрофильные (Arg, Lys, Glu, Asp, Asn, Gln, Ser) или 4 гидрофобные (Phe, Tyr, Trp, Leu, Ile, Val, Met) аминокислоты, он не рассматривался в качестве возможного Т-эпитопа.

В работе [2] отмечается, что некоторые известные Т-эпитопы по распределению гидрофильных и гидрофобных аминокислот ближе к амфипатичным 3_{10} -спиралям, чем к α -спиралям. Эти спиральные структуры различаются по количеству аминокислотных остатков на один виток спирали. Нами были проведены расчеты периодичности в чередовании гидрофильных и гидрофобных остатков, соответствующей как α -спирали (4 аминокислотных остатка на виток, период 90°), так и 3_{10} -спирали (3 остатка на виток, период 120°). Мы использовали период 90° для α -спирали, так как при использовании «канонического» значения 100° были получены худшие результаты. Проверка алгоритма проводилась на белках с известными Т-эпитопами [2]: миоглобине кашалота, цитохромах голубя и быка, свином инсулине, курином лизоциме, курином овальбумине, гематоглиутинине гриппа A/PR8/34, белка VP1 вируса ящура, главном поверхностном антигене вируса гепатита В, белке С1 λ -репрессора, гликопротеине вируса бешенства.

При введении порогового значения алгоритм отбирал участки с абсолютным значением коэффициента корреляции, равным или большим порогового значения (для α -спиралей и 3_{10} -спиралей), идентифицируя их как возможные Т-эпитопы. При пороговом значении, равном 0,6, всего было отобрано 90 таких участков, при этом идентифицировались 17 из 22 известных Т-эпитопов. Для сравнения алгоритм АМРН1 идентифицировал 17 из 22 возможных Т-эпитопов, но при этом было отобрано 111 участков.

Таким образом, при значительно меньшем времени счета и требуемой памяти программа, реализующая разработанный алгоритм, характеризуется достаточно высокой предсказательной способностью. Одно из преимуществ данного подхода заключается в возможности расчета коэффициента корреляции с аperiodическими функциями, которые, возможно, лучше отвечают распределению гидрофильных и гидрофобных аминокислот в Т-эпитопах.

Программа, использующая предложенный алгоритм, реализована на ЭВМ Apple IIe и входит в пакет прикладных иммунохимических программ, разработанных авторами.

СПИСОК ЛИТЕРАТУРЫ

1. DeLisi C., Berzofsky J. // Proc. Nat. Acad. Sci. USA. 1985. V. 82. № 20. P. 7048—7058.
2. Margalit H., Spouge J. L., Cornette J., Cease K., DeLisi C., Berzofsky J. // J. Immunol. 1987. V. 138. № 7. P. 2213—2229.
3. Sette A., Dorio G., Adorini L. // Mol. Immunol. 1986. V. 23. № 8. P. 807—810.
4. Stille C., Thomas L., Reyes V., Humphreys R. // Mol. Immunol. 1987. V. 24. № 10. P. 1021—1027.
5. Дьяконов В. П. Справочник по алгоритмам и программам на языке Бейсик для персональных ЭВМ. М.: Наука, 1987. С. 137.
6. Parker J., Guo D., Hodges R. // Biochemistry. 1986. V. 25. № 19. P. 5425—5432.

После доработки
25.X.1988

Поступило в редакцию
14.VII.1988

A METHOD ON SEARCHING OF AMPHIPATHIC HELICES IN PROTEIN SEQUENCES

GABRIELIAN A. E., IVANOV V. S., KOZHICH A. T.
M. M. Shemyakin Institute of Bioorganic Chemistry, Academy,
of Sciences of the USSR, Moscow

A simple method for searching amphipathic helices based on estimation of correlation between hydrophobicity distribution and periodic function is proposed. The method was examined in series of proteins with known T-cell epitopes, which are mostly amphipathic helices. The predictive power of the method is discussed.