

4. Investigation of restriction-modification enzymes from *M. varians* RFL19 with a new type of specificity toward modification of substrate / V. Butkus, S. Klimasauskas, D. Kersulyte et al. // Nucl. Acids Res.—1985.—13, N 16.— P. 5727—5746.
5. Laemmli U. K. Cleavage of structure proteins during the assembly of the head bacteriophage T4 // Nature.—1970.—227, N 5259.— P. 680—685.
6. Siegel L. M., Monty K. G. Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel-filtration and density gradient centrifugation // Biochim. et biophys. acta.—1966.—112, N 2.— P. 346—362.
7. Herman G. E., Modrich P. *Escherichia coli* dam methylase. Physical and catalytic properties of the homogeneous enzyme // J. Biol. Chem.—1982.—257, N 5.— P. 2605—2612.
8. Martin R. G., Ames B. N. A method for determining the sedimentation behavior of enzymes: application to protein mixtures // Ibid.—1961.—236, N 5.— P. 1372—1379.
9. Януляйтис А. А., Вайткявичюс Д. П. Новый методический подход к разработке получения рестриционных эндонуклеаз. Разработка схемы выделения гомогенного препарата рестриктазы *MvaI* // Биотехнология.—1985.— № 1.— С. 39—51.
10. Alves J. Mechanistische Untersuchungen zur Spaltung von Oligodesoxynucleotiden durch die Restriktionsendonuklease *EcoRI*.—Hannover: Universitat Hannover, 1984.— S. 16—24.
11. Frankel A. D., Ackers G. K., Smith H. O. Measurement of DNA-protein equilibria using gel chromatography: application to the *Hinfi* restriction endonuclease // Biochemistry.—1985.—24, N 12.— P. 3049—3054.

ВНИИ молекуляр. биологии,  
пос. Кольцово Новосиб. обл.  
НПО «Фермент», Вильнюс

Получено 14.07.87

УДК 577.112.855

## ПОДОБИЕ ДНК-УЗНАЮЩИХ СТРУКТУР АКТИВАТОРА КООРДИНИРОВАННОЙ ТРАНСКРИПЦИИ ГЕНОВ, КОДИРУЮЩИХ ФЕРМЕНТЫ БИОСИНТЕЗА АМИНОКИСЛОТ У ДРОЖЖЕЙ, РЕГУЛЯТОРОВ ДИФФЕРЕНЦИРОВКИ КЛЕТОК ДРОЖЖЕЙ И РЕГУЛЯТОРОВ РАЗВИТИЯ И МОРФОГЕНЕЗА

**Б. В. Шестопалов**

**Введение.** У дрожжей *Saccharomyces cerevisiae* существует координация транскрипции генов, кодирующих ферменты биосинтеза аминокислот [1]. Непосредственным активатором координированной транскрипции является белок *GCN4* [2], причем активация связана с узнаванием белком TGAATC последовательности в регуляторных участках активируемых генов [3]. Авторы работы [4] предприняли попытку определить, к какому из двух известных видов ДНК-узнающей структуры, «пальцевидной» или спираль—изгиб—спираль, относится ДНК-узнающая структура белка *GCN4* и где она локализована. В результате выяснено следующее: ДНК-узнающая структура заключена в С-концевом домене белка — сегменте 222—281; «пальцевидная» структура в ДНК-узнающем домене невозможна, так как в его аминокислотной последовательности нет ключевых для нее остатков; присутствие структуры спираль—изгиб—спираль, локализованной по данным предсказания вторичной структуры ДНК-узнающего домена (спираль 244—261, изгиб 262—265, спираль 266—278), сомнительно: спирали длиннее обычных, аминокислотная последовательность не гомологична аминокислотным последовательностям структур спираль—изгиб—спираль. В итоге авторы пришли к выводу, что вопрос о виде ДНК-узнающей структуры белка *GCN4* и ее локализации остается открытым.

В настоящей работе представлены данные, на основании которых сделано заключение, что белок *GCN4*, возможно, узнает ДНК с помощью ДНК-узнающей структуры спираль—изгиб—спираль, локализованной в сегменте аминокислотной последовательности 256—278, и ДНК-узнающая структура белка *GCN4* подобна ДНК-узнающим структурам белков *MATa1* и *MATa2* — регуляторов дифференцировки клеток дрожжей, и гомеодоменсодержащих белков — регуляторов развития и морфогенеза.

**Методы.** Использовали необходимые стереохимические условия существования ДНК-узнающей структуры спираль—изгиб—спираль, описанные в нашей статье [5]

(в дальнейшем — необходимые условия существования). Для сворачивания в эту структуру сегменту аминокислотной последовательности необходимо обладать следующими свойствами: 1) длина сегмента — 23 остатка; 2) остаток в позиции 11 (с N-конца) не должен входить в группу Trp, Tyr, Phe, Ile, Leu, Val, Thr, Pro; остаток в позиции 13 — в группу Trp, Phe, Ile, Leu, Val, Pro; 3) остаток в позиции 17 должен быть из группы Cys, Ile, Val, Leu, Gly, Phe, Ala, Met; остатки в позициях 6 и 10 — из группы, включающей, кроме указанных остатков, также Trp, Ser, Thr, His, Tyr; остаток в позиции 7 — из группы Gly, Ala, Lys, Arg, Cys, Met, Ser; 4) отсутствие Pro в позициях 5—13 и 17—23; 5) в случае существования аминокислотных последовательностей, гомологичных рассматриваемой, их сегменты, соответствующие вышеуказанному сегменту данной последовательности (одинаково с ним расположенные на схеме, где сравниваются все эти последовательности), как правило, также должны обладать приведенными свойствами.

Степень гомологии сегментов аминокислотных последовательностей определяли по числу эквивалентных позиций сегментов, занимаемых идентичными или структурно эквивалентными аминокислотными остатками.

**Результаты и обсуждение.** В ДНК-узнающем домене белка *GCN4* есть только один сегмент, удовлетворяющий необходимым условиям существования ДНК-узнающей структуры спираль—изгиб—спираль [5] — 256—278. Вероятность того, что структура не локализована в этом сегменте, равна 2 % [5]. Предполагаемая локализация ДНК-узнающей структуры согласуется с экспериментальными данными, по которым отщепление C-концевого участка 242—281 приводит к утрате белком способности узнавать ДНК [2]. Ниже приведены другие данные в пользу такой локализации.

Среди известных ДНК-узнающих структур спираль—изгиб—спираль, наблюдаемых или тех из предсказанных, которые отвечают необходимым условиям существования [5], наиболее гомологична предполагаемой ДНК-узнающей структуре белка *GCN4* структура *MATα2* белка дрожжей: в шести эквивалентных позициях сегментов аминокислотных последовательностей находятся идентичные остатки; в одной — структурно-подобные остатки, что дает 30 % гомологии (схема: в [6] цитируется поправка к аминокислотной последовательности белка *MATα1* из [9]; в сплошных (прерывистых)

		Спираль 1	Изгиб	Спираль 2
<i>GCN4</i>	256-278	K V E E L L S K W Y H I E N L Y A R L K K I V		
<i>MATα2</i>	160-182	K G L E N L M K N T S L S R I Q I K N W V S N		
<i>MATα1</i>	99-121	K E K E E Y A K H C G T T P L Q V R V W F I N		
<i>inv</i>	29-51	K R R Q Q L S G E L G L N E A Q I K I W F Q N		

рамках — идентичные (структурно-подобные) остатки в эквивалентных позициях структур спираль—изгиб—спираль двух или нескольких белков, включая белок *GCN4*; сплошные (прерывистые) подчеркивания — то же, исключая белок *GCN4*; структура спираль—изгиб—спираль белка, кодируемого геном *inv* дрозофилы, выбрана из структур спираль—изгиб—спираль, предполагаемых в гомеодоменсодержащих белках, из-за того, что ее аминокислотная последовательность наиболее гомологична таковой предполагаемой структуры спираль—изгиб—спираль белка *GCN4*). Подобная степень гомологии обнаружена между предполагаемыми ДНК-узнающими структурами спираль—изгиб—спираль белков *MATα2* и *MATα1* и гомеодоменсодержащих белков [6, 7]. Это навело нас на мысль объединить белки *GCN4*, *MATα2* и *MATα1* в гомеодоменсодержащие в группу белков, гомологичных по ДНК-узнающей структуре. На схеме представлены данные в пользу такого объединения: в семи эквивалентных позициях сегментов этих белков, содержащих сравнимые структуры, находятся идентичные или структурно-подобные аминокислотные остатки (для трех из четырех белков, включая белок *GCN4*). Очевидно, что в целом по группе гомология сосредоточена в области спирали 1 и практически отсутствует в области спирали 2. Поскольку, как известно, спираль 2 участвует в узнавании белком ДНК, а спираль 1 — не участвует, слабая гомология структур в области спирали 2 может быть отражением слабой гомологии узнаваемых этими структурами последовательностей ДНК.

К данным в пользу объединения этих белков в одну группу, кроме уже приведенных в настоящей работе или содержащихся в работах [6, 9, 11], можно также отнести следующие: 1) в гомеодоменах, где, как показано на примере белка, кодируемого геном *inv* дрозофилы [11], заключена ДНК-узнающая активность всего белка, предполагаемая локализация ДНК-узнающей структуры спираль—изгиб—спираль является единственной, отвечающей необходимым условиям существования [5], причем условия существования второй из двух известных ДНК-узнающих структур, «назвече-

видной», в гомеодомене не выполняются; 2) идентифицированная в [6] температурочувствительная мутация в гене *fitz*, *fitz*<sup>1471s</sup> связана с аминокислотной заменой, запрещенной необходимыми условиями существования ДНК-узнающей структуры спираль—изгиб—спираль [5, 9] (Ala на Val в седьмой с N-конца позиции сегмента 256—278).

Из литературных данных и данных настоящей работы следует, что подобные друг другу структуры спираль—изгиб—спираль белков *GCN4*, *MATα2* и *MATα1* и гомеодоменосодержащих белков, вероятно, необходимы для узнавания ДНК, что в свою очередь необходимо для осуществления ими соответствующих регуляторных функций в столь различных биологических процессах, как внутриклеточный метаболизм, дифференцировка клеток, развитие и морфогенез многоклеточных организмов [2, 11—14].

#### SIMILARITY OF DNA-RECOGNIZING STRUCTURES OF THE COORDINATED TRANSCRIPTION ACTIVATOR OF YEAST AMINO ACID BIOSYNTHESIS ENZYMES GENES, YEAST CELLS DIFFERENTIATION REGULATORS, DEVELOPMENT AND MORPHOGENESIS REGULATORS

B. V. Shestopalov

Institute of Cytology, Academy of Sciences of the USSR, Leningrad

#### Summary

Based on the data obtained by the method of necessary stereochemical requirements of the DNA-recognizing structure helix-turn-helix existence and the available data from literature a conclusion is made that 1) *GCN4* protein, probably, recognizes DNA using DNA-recognizing structure helix-turn-helix, localized in the amino acid sequence segment 256-278 and 2) DNA-recognizing structure of *GCN4* protein is similar to those of *MATα1* and *MATα2* proteins, and homeodomain-containing proteins.

1. Hinnebusch A. G. Multiple level of gene regulations in the control of amino acid biosynthesis in *Saccharomyces cerevisiae* // BioEssays.— 1986.— 5, N 2.— P. 57—62.
2. Hope I. A., Struhl K. *GCN4* protein, synthesized in vitro, binds *HIS3* regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast // Cell.— 1985.— 43, N 1.— P. 177—188.
3. Saturation mutagenesis of the yeast *his3* regulatory site: requirements for transcriptional induction and for binding by *GCN4* activator protein / D. E. Hill, I. A. Hope, J. P. Macke, K. Struhl // Science.— 1986.— 234, N 4775.— P. 451—457.
4. Hope I. A., Struhl K. Functional dissection of a eukaryotic transcriptional activator protein, *GCN4* of yeast // Cell.— 1986.— 46, N 6.— P. 885—894.
5. Шестопалов Б. В. Предсказание ДНК-узнающей супервторичной структуры белка спираль—изгиб—спираль по методу модифицированных необходимых стереохимических условий существования Олендорфа—Андерсона—Метьюза // Молекуляр. биология.— 1988.— 22, № 2.— С. 923—930.
6. Laughon A., Scott M. P. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins // Nature.— 1984.— 310, N 5972.— P. 25—31.
7. Fly and frog homeo domains show homologies with yeast mating type regulatory proteins / J. C. W. Shepherd, W. McGinnis, A. E. Corrasco et al. // Ibid.— P. 70—71.
8. Porter S. D., Smith M. Homoco-domain homology in yeast *MATα2* is essential for repressor activity // Ibid.— 1986.— 320, N 6064.— P. 766—768.
9. Ohlendorf D. H., Anderson W. F., Matthews B. W. Many gene-regulatory proteins appear to have a similar α-helical fold that binds DNA and evolved from a common precursor // J. Mol. Evol.— 1983.— 19, N 1.— P. 109—114.
10. The engrailed locus of *Drosophila*: structural analysis of an embryonic transcript / S. J. Poole, L. M. Kauvar, B. Drees, T. Kornberg // Cell.— 1985.— 40, N 1.— P. 37—43.
11. Desplan C., Theis J., O'Farrell P. H. The *Drosophila* development gene, engrailed, encodes a sequence specific DNA-binding activity // Nature.— 1985.— 318, N 6047.— P. 630—635.
12. Johnson A. D., Herskowitz I. A repressor (*MATα2* product) and its operator control expression of a set of cell type specific genes in yeast // Cell.— 1985.— 42, N 1.— P. 237—247.
13. Miller A. M., Mackay V. L., Nasmyth K. A. Identification and comparison of two sequence elements that confer cell-type specific transcription in yeast // Nature.— 1985.— 314, N 6012.— P. 598—603.
14. The homeo domain of a murine protein binds 5' to its own homeo box / A. Fainsod, L. D. Bogorad, T. Ruusala et al. // Proc. Nat. Acad. Sci. USA.— 1986.— 83, N 24.

Ин-т цитологии АН СССР, Ленинград  
P. 9532—9536.

Получено 14.07.87