***Белок биотинового биосинтеза BIOH\_ECOLI***

**Белок BIOH\_ECOLI** есть не что иное, как карбоксилэстераза bioH, или белок биотинового синтеза bioH (принадлежит к надсемейству АВ гидролаз, семейству карбоксилэстераз bioH); найден в бактерии Escherichia coli (strain K12) (Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia). Фермент проявляет себя как карбоксилэстераза преимущественно для коротких (до 6 углеродов) углеводородных цепочек эфиров жирных кислот. Кроме того, белок может выполнять функцию тиоэстеразы. Карбоксилэстераза bioH может образовывать комплекс с CoA и участвовать в реакции между СоА и пимелиновой кислотой с образованием пимелил-СоА, предшественника биотина в биосинтезе.

**Биоти́н** (витамин Н1, витамин B7, коэнзим R) — водорастворимый витамин, входящий в состав ферментов, регулирующих белковый и жировой обмен. Он участвует в синтезе глюкокиназы (фермента, катализирующего фосфорилирование шестиатомных сахаров (гексоз)), а также является коферментом различных ферментов. С участием биотина протекают реакции активирования, переносаи фиксации СО2(взято из Википедии - <http://www.wikipedia.org/>). Таким образом, биотин является жизненно важным веществом для многих организмов, некоторые из которых могут синтезировать его сами, как, например, зеленые растения, некоторые грибы и бактерии (в том числе и E. coli).

Рассмотрим особенности белка биотинового биосинтеза бактерии E. coli.

Энзим располагается вероятнее всего в цитоплазме клетки и функционирует наилучшим образом при рН, равном 8.0-8.5.

Белок состоит из одной субъединицы - белковой глобулы, образованной одной полипептидной цепочкой из 256 аминокислот (<http://kodomo.cmm.msu.ru/~lu.andreeva/bioh_fasta.html>). Активный центр образуют 82-я (серин), 207-я (аспарагиновая кислота) и 235-я (гистидин) аминокислоты. Возможна мутация 82 аминокислоты, не влияющая на связывание CoA, при этом серин замещается аланином. Средствами масс спектрометрии была установлена молекулярная масса белка, равная 28502.0 А.е. с погрешностью 5.5.

Рассмотрим третичную структуру белка (рисунок сделан средствами программы RasMol на основе данных документа 1M33.pdb банка PDB (<http://www.pdb.org/pdb/files/1m33.pdb>)):



На рисунках розовым цветом обозначены альфа-спирали, жёлтым – бета-тяжи, а серым – участки нерегулярной структуры.

Из информации, предложенной в документе UniProt (<http://www.uniprot.org/uniprot/P13001.txt>), и из приведённых выше рисунков можно сделать вывод, что белок в равной степени состоит из альфа-спиралей и бета-тяжей, чередующихся друг с другом. Это подтверждается и распределением торсионных углов фи и пси данного белка (<http://kodomo.cmm.msu.ru/~lu.andreeva/torsion.html>).

 Белок BIOH\_ECOLI нельзя назвать очень популярным объектом исследований в научном сообществе: с помощью поисковой системы SRS (<http://srs.ebi.ac.uk/>) было найдено 55-60 статей с упоминанием данного энзима, а в известной базе данных PubMed (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed>) обнаружилось только 9 статей. Однако в документе базы данных UniProt (<http://www.uniprot.org/uniprot/P13001.txt>) представлено 5 статей о карбоксилэстеразе bioH, публиковавшихся с 1989 по 2006 год, что говорит о том, что белок всё-таки исследуется.

Аннотации и ссылки на статьи, упомянутые в документе UniProt (<http://www.uniprot.org/uniprot/P13001.txt>) представлены ниже, полного текста статей нет в открытом доступе.

1. **"Nucleotide sequence of the bioH gene of Escherichia coli"**

O'Regan M., Gloeckler R., Bernard S., Ledoux C., Ohsawa I., Lemoine Y.

 Nucleic Acids Res. 17:8004-8004(1989).

 <http://www.ncbi.nlm.nih.gov/pubmed/2678009?ordinalpos=5&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_DefaultReportPanel.Pubmed_RVDocSum>

1. **“The complete genome sequence of Escherichia coli K-12”**

Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.

Science. 1997 Sep 5;277(5331):1453-74.

Laboratory of Genetics, University of Wisconsin-Madison, 445 Henry Mall, Madison, WI 53706, USA. ecoli@genetics.wisc.edu

The 4,639,221-base pair sequence of Escherichia coli K-12 is presented. Of 4288 protein-coding genes annotated, 38 percent have no attributed function. Comparison with five other sequenced microbes reveals ubiquitous as well as narrowly distributed gene families; many families of similar genes within E. coli are also evident. The largest family of paralogous proteins contains 80 ABC transporters. The genome as a whole is strikingly organized with respect to the local direction of replication; guanines, oligonucleotides possibly related to replication and recombination, and most genes are so oriented. The genome also contains insertion sequence (IS) elements, phage remnants, and many other patches of unusual composition indicating genome plasticity through horizontal transfer.

<http://www.ncbi.nlm.nih.gov/pubmed/9278503?ordinalpos=1&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_DefaultReportPanel.Pubmed_RVDocSum>

1. **“Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110”**

Hayashi K., Morooka N., Yamamoto Y., Fujita K., Isono K., Choi S., Ohtsubo E., Baba T., Wanner B.L., Mori H., Horiuchi T.

Mol Syst Biol. 2006;2:2006.0007. Epub 2006 Feb 21

Division of Genome Dynamics, National Institute for Basic Biology, Myodaiji, Okazaki, Aichi Pref., Japan.

With the goal of solving the whole-cell problem with Escherichia coli K-12 as a model cell, highly accurate genomes were determined for two closely related K-12 strains, MG1655 and W3110. Completion of the W3110 genome and comparison with the MG1655 genome revealed differences at 267 sites, including 251 sites with short, mostly single-nucleotide, insertions or deletions (indels) or base substitutions (totaling 358 nucleotides), in addition to 13 sites with an insertion sequence element or defective prophage in only one strain and two sites for the W3110 inversion. Direct DNA sequencing of PCR products for the 251 regions with short indel and base disparities revealed that only eight sites are true differences. The other 243 discrepancies were due to errors in the original MG1655 sequence, including 79 frameshifts, one amino-acid residue deletion, five amino-acid residue insertions, 73 missense, and 17 silent changes within coding regions. Errors in the original MG1655 sequence (<1 per 13,000 bases) were mostly within portions sequenced with out-dated technology based on radioactive chemistry.

<http://www.ncbi.nlm.nih.gov/pubmed/16738553?ordinalpos=2&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_DefaultReportPanel.Pubmed_RVDocSum>

1. **“Purification and characterisation of the BIOH protein from the biotin biosynthetic pathway”**

Tomczyk N.H., Nettleship J.E., Baxter R.L., Crichton H.J., Webster S.P., Campopiano D.J.

FEBS Lett. 2002 Feb 27;513(2-3):299-304

Department of Chemistry, Joseph Black Building, University of Edinburgh, West Mains Road, EH9 3JJ, Edinburgh, UK.

Conversion of pimeloyl-coenzyme A (CoA) to biotin in Escherichia coli requires at least four enzymes encoded by genes in the bio operon. One gene, bioH, which is not present in the bioABFCD operon, is required for the synthesis of pimeloyl-CoA but its exact role in formation of this intermediate is unknown. To investigate this further, we have overexpressed and purified the bioH gene products from both E. coli (BIOH EC) and Neisseria meningitis (BIOH NM) in E. coli. When purified BIOH was incubated with excess CoA and analysed by electrospray mass spectrometry a species of mass corresponding to a BIOH:CoA complex was observed. Mutation of a conserved serine residue to alanine (BIOH EC S82A) did not prevent CoA binding. This is the first report of the purification of BIOH and the observation of a small molecule bound to the protein provides clues to its role in pimeloyl-CoA synthesis.

<http://www.ncbi.nlm.nih.gov/pubmed/11904168?ordinalpos=1&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_DefaultReportPanel.Pubmed_RVDocSum>

1. **“Integrating structure, bioinformatics, and enzymology to discover function: BioH, a new carboxylesterase from Escherichia coli”**

Sanishvili R., Yakunin A.F., Laskowski R.A., Skarina T., Evdokimova E., Doherty-Kirby A., Lajoie G.A., Thornton J.M., Arrowsmith C.H., Savchenko A., Joachimiak A., Edwards A.M.

J Biol Chem. 2003 Jul 11;278(28):26039-45. Epub 2003 May 5

Biosciences Division, Argonne National Laboratory, Argonne, Illinois, 60439, USA.

Structural proteomics projects are generating three-dimensional structures of novel, uncharacterized proteins at an increasing rate. However, structure alone is often insufficient to deduce the specific biochemical function of a protein. Here we determined the function for a protein using a strategy that integrates structural and bioinformatics data with parallel experimental screening for enzymatic activity. BioH is involved in biotin biosynthesis in Escherichia coli and had no previously known biochemical function. The crystal structure of BioH was determined at 1.7 A resolution. An automated procedure was used to compare the structure of BioH with structural templates from a variety of different enzyme active sites. This screen identified a catalytic triad (Ser82, His235, and Asp207) with a configuration similar to that of the catalytic triad of hydrolases. Analysis of BioH with a panel of hydrolase assays revealed a carboxylesterase activity with a preference for short acyl chain substrates. The combined use of structural bioinformatics with experimental screens for detecting enzyme activity could greatly enhance the rate at which function is determined from structure.

<http://www.ncbi.nlm.nih.gov/pubmed/12732651?ordinalpos=1&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_DefaultReportPanel.Pubmed_RVDocSum>

Все вышеперечисленные статьи найдены в базе данных PubMed (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed>). Названия и авторов других статей в PubMed с упоминанием белка можно найти здесь: <http://kodomo.cmm.msu.ru/~lu.andreeva/pubmed.html>.