Критический анализ модели белка инозозоизомеразы *Вacillus Subtilis*, представленном в банке PDB (1I6N).

Отчет Караваевой Ю.Е., Факультет Биоинженерии и Биоинформатики, МГУ 2012.

В отчете приведены результаты анализа качества структуры инозозоизомеразы, расшифрованной методом РСА авторами [R.-G. Zhang](http://www.ncbi.nlm.nih.gov/pubmed/?term=Zhang%20RG%5Bauth%5D), [I. Dementieva](http://www.ncbi.nlm.nih.gov/pubmed/?term=Dementieva%20I%5Bauth%5D), [N. Duke](http://www.ncbi.nlm.nih.gov/pubmed/?term=Duke%20N%5Bauth%5D), [F. Collart](http://www.ncbi.nlm.nih.gov/pubmed/?term=Collart%20F%5Bauth%5D), [E. Quaite-Randall](http://www.ncbi.nlm.nih.gov/pubmed/?term=Quaite-Randall%20E%5Bauth%5D), [R. Alkire](http://www.ncbi.nlm.nih.gov/pubmed/?term=Alkire%20R%5Bauth%5D), [L. Dieckman](http://www.ncbi.nlm.nih.gov/pubmed/?term=Dieckman%20L%5Bauth%5D), [N. Maltsev](http://www.ncbi.nlm.nih.gov/pubmed/?term=Maltsev%20N%5Bauth%5D), [O. Korolev](http://www.ncbi.nlm.nih.gov/pubmed/?term=Korolev%20O%5Bauth%5D), и [A. Joachimiak](http://www.ncbi.nlm.nih.gov/pubmed/?term=Joachimiak%20A%5Bauth%5D) в 2002 году (журнал Proteins). Код структуры в банке PDB: 1I6N.

1. **Введение.**

Мио-инозитол используется почвенными микроорганизмами как источник углерода. Гены, участвующие в метаболизме мио-инозитола плохо охарактеризованы. В геноме B. subtilis путь катаболизма инозитола состоит из нескольких генов, в том числе ген IOLI (NCBI P42419). Этот ген кодирует консервативный белок, состоящий из 278 аминокислотных остатков, с неизвестной функцией. Предполагалось, что этот белок может обладать неизвестной ранее пространственной структурой.

Методом РСА была получена кристаллическая структура этого белка с разрешением 1.6 Å с использованием метода полуватоматического высокопроизводительного подхода. Анализ кристаллической структуры 1I6N показал, что белок обладает структурой тим-баррель.

1. **Результаты.**

# Статья «Crystal Structure of *Bacillus subtilis* IolI Shows Endonuclase IV Fold With Altered Zn Binding» авторов [*R.-G. Zhang*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Zhang%20RG%5Bauth%5D)*,* [*I. Dementieva*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Dementieva%20I%5Bauth%5D)*,* [*N. Duke*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Duke%20N%5Bauth%5D)*,* [*F. Collart*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Collart%20F%5Bauth%5D)*,* [*E. Quaite-Randall*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Quaite-Randall%20E%5Bauth%5D)*,* [*R. Alkire*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Alkire%20R%5Bauth%5D)*,* [*L. Dieckman*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Dieckman%20L%5Bauth%5D)*,* [*N. Maltsev*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Maltsev%20N%5Bauth%5D)*,* [*O. Korolev*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Korolev%20O%5Bauth%5D)*, и* [*A. Joachimiak*](http://www.ncbi.nlm.nih.gov/pubmed/?term=Joachimiak%20A%5Bauth%5D) опубликована в журнале Proteins в 2002 году (Proteins. 2002 August 1; 48(2): 423–426).

# Исследованный белок состоит из одной полипептидной цепи, имеющей третичную структуру тим-баррель.

# E:\julia\Учеба\7_семестр\bioinf\prac4\1.png

# Рис.1. Пространственная структура белка 1I6N.

# Для решения фазовой проблемы использован метод многоволнового аномального рассеяния (MAD).

# Максимальное и минимальное разрешение для использованных рефлексов составляет 19.90 - 1.80 Å, разрешение структуры белка в PDB составляет 1.80 Å с полнотой данных 98%, что является неплохим результатом.

# R-фактор составляет 0.209, R-free фактор - 0.238. Такие значения говорят о том, что модель соответствует экспериментальным данным, а разница между R-free и R-фактором, составляющая 0.029 говорит о хорошей оптимизации структуры.

# Кристаллическая ячейка ортогональна, длины базисных периодов ячейки составляют 74.015, 104.616, 48.462 Å. Тип центрирования P 21 21 2.

# Поиск маргинальных остатков.



Рис.2. Карта Рамачандрана для белка 1I6N.

Plot statistics

Residues in most favoured regions [A,B,L] 235 92.5%

Residues in additional allowed regions [a,b,l,p] 17 6.7%

Residues in generously allowed regions [~a,~b,~l,~p] 2 0.8%

Residues in disallowed regions 0 0.0%

Number of non-glycine and non-proline residues 254 100.0%

Number of end-residues (excl. Gly and Pro) 2

Number of glycine residues (shown as triangles) 13

Number of proline residues 8

Total number of residues 277

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms

and R-factor no greater than 20%, a good quality model would be expected

to have over 90% in the most favoured regions

На карте Рамачандрана для всего белка (рис.2) видно, что ни один остаток не попал в запрещенную область, однако два остатка попали в допустимую область и предположительно являются маргиналами (Asp202 и Ala67). Поскольку структура не является переоптимизированной, о чем говорит разница значений R-фактора и R-free, можно предположить, что полученная структура хорошо соответствует настоящей структуре белка. Карты Рамачандрана для каждого типа аминокислотных остатков можно посмотреть в прилагающихся файлах.

На картах торсионных углов ротамеров белка 1I6N потенциальным маргинальным остатком является Ala209 (карты в прилагающихся файлах).

## RSR (Real-space R-value vs Residue for 1i6n)



Рис.3. Значения RSR для белка 1I6N.

Большое значение RSR для остатка говорит о том, что этот остаток имеет плохую плотность по сравнению со средним значением для аналогичных остатков с похожим разрешением. В данном белке довольно много остатков с высокими значениями RSR, и они распределены более-менее равномерно по длине последовательности, что говорит о том, что в структуре нет большого участка, плохо вписанного в электронную плотность. Есть 9 аминокислотных остатков, которые имеют значение RSR значительно выше 0.2. Это основные кандидаты в маргинальные остатки:

254 LYS, Value=0.229
241 ASP, Value=0.315
236 GLU, Value=0.233
194 LYS, Value=0.209
119 LYS, Value=0.235
67 ALA, Value=0.284
56 GLN, Value=0.270
53 GLU, Value=0.224
44 LYS, Value=0.304

Рассмотрим остаток ASP241, обладающий наибольшим значением SRS и ALA67, который также находится в допустимой области на карте Рамачандрана.



 А В

Рис.4. Маргинальные остатки, плохо вписанные в электронную плотность.

 (А) – Ala67, (B) – Asp241. Срез электронной плотности на уровне 1.5.

1. **Заключение.**
2. Значения глобальных факторов говорят о том, что полученная структура неплохо соответствует экспериментальным данным и не является переоптимизированной.
3. Судя по значениям SRS, в структуре есть несколько маргинальных остатков, плохо вписанных в электронную плотность, но число таких остатков не превышает 6% от общего числа остатков.
4. В выдаче программы WHAT\_CHECК говорится много о плохой геометрии многих аминокислотных остатков в последовательности, однако на карте Рамачандрана и торсионных углов для ротамеров не видно особо серьезных нарушений, за исключением Asp202 и Ala67, которые часто фигурируют в ошибках, отмеченных WHAT\_CHECК.
5. Поскольку авторы предполагали, что данный белок обладает ранее не обнаруженной пространственной структурой, они не использовали метод молекулярного замещения и, скорее всего, не использовали структуры гомологичных белков для улучшения данной структуры. В связи с этим можно сказать, что структура расшифрована очень хорошо.
6. **Источники**

# [R.-G. Zhang](http://www.ncbi.nlm.nih.gov/pubmed/?term=Zhang%20RG%5Bauth%5D), [I. Dementieva](http://www.ncbi.nlm.nih.gov/pubmed/?term=Dementieva%20I%5Bauth%5D), [N. Duke](http://www.ncbi.nlm.nih.gov/pubmed/?term=Duke%20N%5Bauth%5D), [F. Collart](http://www.ncbi.nlm.nih.gov/pubmed/?term=Collart%20F%5Bauth%5D), [E. Quaite-Randall](http://www.ncbi.nlm.nih.gov/pubmed/?term=Quaite-Randall%20E%5Bauth%5D), [R. Alkire](http://www.ncbi.nlm.nih.gov/pubmed/?term=Alkire%20R%5Bauth%5D), [L. Dieckman](http://www.ncbi.nlm.nih.gov/pubmed/?term=Dieckman%20L%5Bauth%5D), [N. Maltsev](http://www.ncbi.nlm.nih.gov/pubmed/?term=Maltsev%20N%5Bauth%5D), [O. Korolev](http://www.ncbi.nlm.nih.gov/pubmed/?term=Korolev%20O%5Bauth%5D), и [A. Joachimiak](http://www.ncbi.nlm.nih.gov/pubmed/?term=Joachimiak%20A%5Bauth%5D). Crystal Structure of Bacillus subtilis IolI Shows Endonuclase IV Fold With Altered Zn Binding. Proteins. 2002 August 1; 48(2): 423–426 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2792003/>)

#  <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=1i6n>

# <http://swift.cmbi.ru.nl/gv/pdbreport/>