

# Обзор протеома бактерии

Горелова Ирина Константиновна

---

Факультет биоинженерии и биоинформатики, Московский государственный университет имени М. В. Ломоносова,

Ленинские горы д. 1 стр. 73, 119234, Москва, Россия

---

## Резюме

В данной работе представлен анализ генома и протеома бактерии *Thermoanaerobacterium thermosaccharolyticum* DSM 571.

Данные о количестве генов белков и генов РНК по категориям, предоставленные сайтом NCBI, были проанализированы с помощью MS Excel. По такому же принципу было рассчитано количество генов белков, псевдогенов, и генов РНК на прямой и комплементарной цепочках ДНК.

Были рассмотрены и проанализированы длины белков, составлена статистика на основе результатов анализа, построена гистограмма распределения.

Выполнена оценка числа генов на 1 млн пар нуклеотидов.

Ключевые слова: *Thermoanaerobacterium thermosaccharolyticum* DSM 571, геном, протеом, бактерия, белок.

## Введение

*Thermoanaerobacterium thermosaccharolyticum* DSM 571 (штамм ATCC 7956) - анаэробная, грамположительная, т. е. сохраняет окраску, не обесцвечивается при промывке и при применении окраски микроорганизмов по методу Грама, термофильная бактерия, использует различные ферменты для расщепления крахмала, амилозы и пуллулана. Условия для оптимальной деятельности энзима: температура 65°C и pH 6,5. Эти бактерии используются в качестве эффективного средства преобразования целлюлозных соединений в этанол и другие побочные продукты.

*Thermoanaerobacterium thermosaccharolyticum* DSM 571 ассоциируется с *C. thermocellum* в высокотемпературных, целлюлозно-расщепляющих средах. Она была предложена в качестве эффективного средства преобразования целлюлозных соединений в этанол и другие побочные продукты. Для этого вида была разработана генетическая система.

По вопросу влияния pH на жизнедеятельность бактерии мною было найдено описание исследования.

Заинтересовавшись свойствами термофильных спорообразующих бактерий и возможностью их практического применения в пищевой отрасли, было проведено исследование, целью которого было определение и моделирование влияния pH на рост, термостойкость и восстановительные способности после термообработки *T. thermosaccharolyticum* DSM 571. Результатами этой работы стал вывод о том, что наибольшие тепловые сопротивления спор и наибольшая способность к восстановлению после термообработки получены при pH-состоянии, позволяющем достичь максимальной скорости. Кроме того, эта работа предполагает, что *T. thermosaccharolyticum* сохраняет свою способность к прорастанию после термообработки при низком pH, где другие высокорезистентные спорообразующие бактерии, такие как *Geobacillus stearothermophilus*, не могут расти.

Имя	<i>Thermoanaerobacterium thermosaccharolyticum</i> DSM 571
Инвентарный номер	NC_014410
Таксономия	
Царство:	Бактерии
Тип:	Firmicutes
Класс:	Клостридия
Порядок:	Thermoanaerobacterales
Семья:	Thermoanaerobacterales семья III. Incertae Sedis
Род:	Thermoanaerobacterium
Вид:	Thermoanaerobacterium thermosaccharolyticum
Штамм	DSM 571
Центр секвенирования	(11-авг-2010) Объединенный институт генома, 2800 Mitchell Drive B310, Walnut Creek, CA 94598-1698, США

	(16-авг-2010) Национальный Центр информации о биотехнологии, NIH, Bethesda, MD 20894, США
Метод секвенирования	Illumina, 454
Форма	Бациллы, Стержневой
Мобильность, подвижность	Да
Присутствие флагеллара	Да
Количество мембран	1
Требования к кислороду	Анаэробный
Оптимальная температура	60-65°C
Температурный диапазон	Термофильный
Естественная среда обитания	Специализированный
Биотические отношения	Свободноживущие
Расположение клеток	Пары, Одиночные
Патогенность	Нет
Эндоспоры	Да
Хромосома	Кольцевая

**Таблица 1. Общие сведения о *Thermoanaerobacterium thermosaccharolyticum* DSM 571**

## Материалы и методы

Геном был скачан с сайта NCBI из банка GenBank.

<https://www.ncbi.nlm.nih.gov/genome/genomes/1864>

Для обработки информации, предоставленной сайтом NCBI, был использован MS Office Excel 2016.

Скаченный файл можно посмотреть по ссылке:

[ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/145/615/GCA\\_000145615.1\\_ASM14561v1](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/145/615/GCA_000145615.1_ASM14561v1)

## Результаты и обсуждение

## Геном

Геном бактерии состоит из генов, кодирующих белки, псевдогенов, нефункциональных аналогов структурных генов и генов, кодирующих РНК, расположенных на прямой и комплементарной цепочках ДНК. Из *Таблицы 2* видно, что большинство генов кодируют белок (~94%).

Strand	Protein	Pseudogene	RNA
Positive	505	23	11
Negative	1996	79	56
Total	2501	102	67

**Таблица 2. Число генов белков, псевдогенов и генов РНК на прямой и комплементарной цепи**

Среди генов, кодирующих белки выделяют категории, распределение по которым также неравномерно: гипотетические, рибосомальные, транспортные и другие. В частности, наибольшее число генов, кодирующих белки, приходится на белки из категории “Другие” (other) (~66%), за ними следуют гипотетические белки (hypothetical) (~23%), потом транспортные (transport) (~9%) и меньше всего генов кодирует рибосомальные (ribosomal) (~2%) (*Таблица 3*) Большое содержание белков из категорий “гипотетические” и “другие” говорит о том, что геном этого штамма плохо изучен.

Strand	Hypothetical	Ribosomal	Transport	Other
Positive	128	7	54	316
Negative	445	56	165	1330
Total	573	63	219	1646

**Таблица 3. Число генов белков по категориям: гипотетические, рибосомальные, транспортные, другие**

Генов, кодирующих РНК гораздо меньше, чем генов, кодирующих белки: на 1 ген РНК приходится примерно 39 генов белков. РНК подразделяются на классы: рибосомальные, транспортные, транспортно-матричные и ядерные. Из *Таблицы 4* видно, что подавляющее число генов кодируют тРНК (~85%).

Strand	rRNA	tRNA	tmRNA	ncRNA	Total
Positive	0	11	0	0	11
Negative	5	44	1	4	54
Total	5	55	1	4	65

**Таблица 4. Число генов РНК по категориям: рибосомальные, транспортные, ядерные**

Также нетрудно заметить, что по, по большей части, геном бактерии закодирован комплементарной цепочкой.

## Протеом

Протеом *Thermoanaerobacterium thermosaccharolyticum* DSM 571 состоит из 2501 белка, среди которых присутствуют транспортные, гипотетические, рибосомальные и другие.

Длина белков бактерии варьирует от 37ак до 2874ак, что согласуется с довольно сильным стандартным отклонением, составляющим ~201ак (Таблица 5). Этот факт говорит о большом разнообразии белков, входящих в структуру бактерии. Как показывает Гистограмма 1, подавляющее большинство белков имеют длину в районе 200-300ак, в пределах этой длины находится медиана, а среднее значение лишь ненамного выходит из этого диапазона. За пределами этой области число белков постепенно падает.

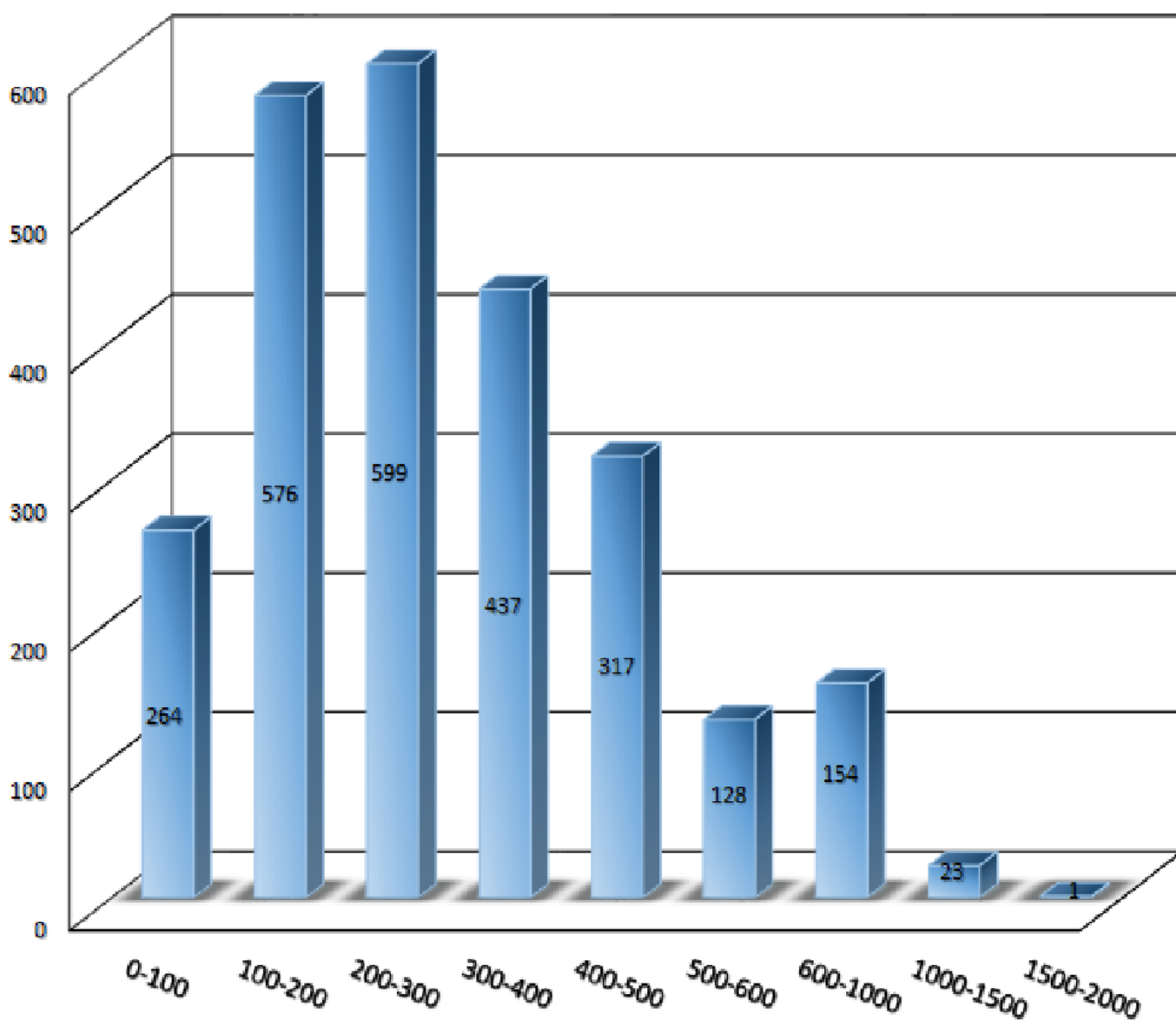
Минимальная длина	37
Максимальная длина	2874
Средняя длина	305.4052
Стандартное отклонение	201.1524923
Медиана	268

Таблица 5. Статистика длин белков

В гистограмму не вошёл самый длинный белок, имеющий длину 2874ак. Значение его длины существенно оторвано от предыдущего (1753ак). Его наличие объясняет разницу между медианой и средним значением. Этот белок – гликозилтрансфераза [1] (glycosyl transferase) (Рисунок 1) – фермент, переносящий остатки моносахаридов от углевода-донора на молекулу-акцептор, чаще всего на спирт. Продуктами таких реакций могут быть моносахариды, гликозиды, олигосахариды, полисахариды, а также гликопротеиды. Некоторые гликозилтрансферазы могут также переносить остатки неорганического фосфата или воду.

Самым коротким белком является рибосомальный белок (50S ribosomal protein L36)

Length of proteins (*Thermoanaerobacterium thermosaccharolyticum* DSM 571)



**Гистограмма 1. Распределение белков по их длинам, которая позволяет оценить количество белков, имеющих длины от 0 до 2000 сначала с шагом в 100 аминокислот, а после 600 в 500 аминокислот**

Оценка числа генов на 1 млн пар нуклеотидов: 958,45.

# Сопроводительные материалы

[..\Documents\Excel\GCA\\_002701205.1\\_ASM270120v1\\_feature\\_table.xlsx](..\Documents\Excel\GCA_002701205.1_ASM270120v1_feature_table.xlsx)

## Благодарности

Я хочу поблагодарить Салимгареева Руслана, Моргуляна Павла, Барсукову Анастасию за помощь в составлении обзора.

## Список литературы

[1] Glycosyltransferase: [Электронный ресурс]. – Режим доступа:

<https://en.wikipedia.org/wiki/Glycosyltransferase> - Дата обращения: 14.03.2019.

<https://string-db.org>

<https://genome.jgi.doe.gov/portal/theth/theth.home.html>

<http://www.straininfo.net/genomes/33165>

[https://www.genome.jp/dbget-bin/www\\_bget?gn:T01299](https://www.genome.jp/dbget-bin/www_bget?gn:T01299)

<https://www.semanticscholar.org/paper/Effect-of-pH-on-Thermoanaerobacterium-DSM-571-spore-Mtimet-Guégan/da7cbb2060e06541edb8294dc891afdbbd55b8d0>