
The proteome of *Denitrovibrio acetiphilus* DSM 12809

Short review

Kozlova Anastasiya

The faculty of Bioengineering and
Bioinformatics, MSU

Motivation: The main objective of the work was a study of the proteome of *Denitrovibrio acetiphilus* DSM 12809, the bacterium. The central subjects of the study were: the distribution of proteins based on their length and the variations of location of genes on different DNA strands.

1. Introduction

Denitrovibrio acetiphilus DSM 12809 is a vibrio shaped gram- anaerobe bacterium, special for the lack of plasmids in its genome. The whole genome contains 3069 genes, wherein 2964 of them encode proteins and others – t-and rRNA. There are no diseases, which might be caused by this organism. The bacterium is free living and typically inhabits fresh water as well as marine.

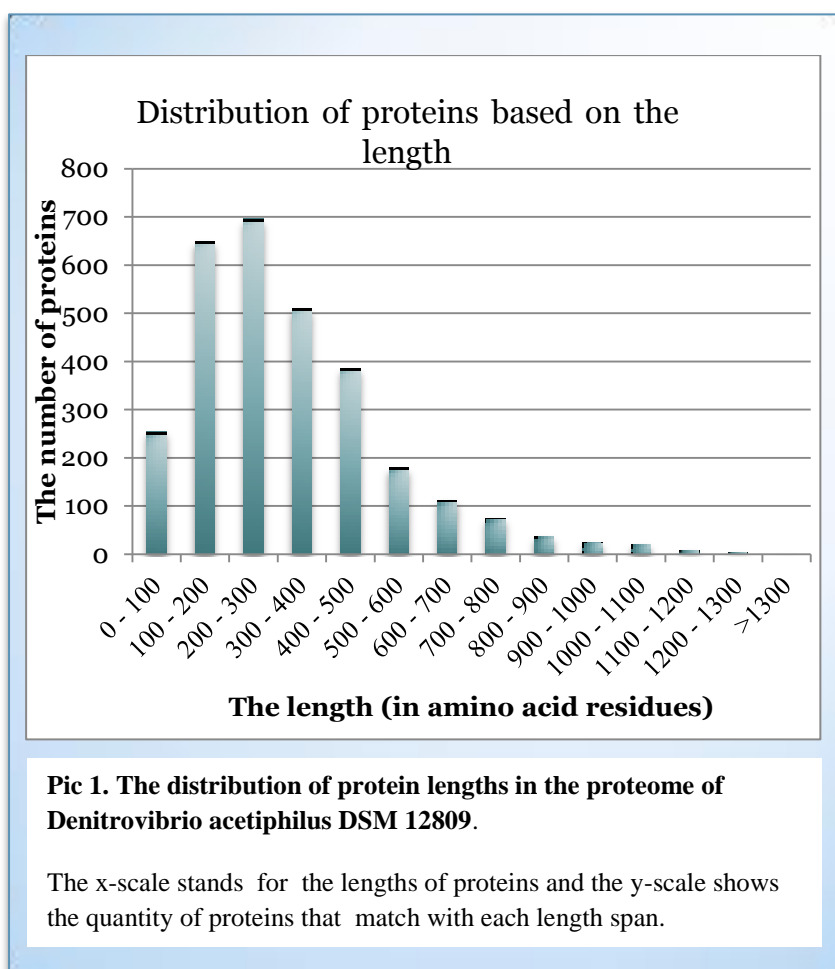
Our work is designed to find out the quantity of encoded proteins in each of the span of their lengths. Moreover, during the study, the number of genes in both straight and complement DNA strands was set up.

2. Methods

The information about the bacterium was taken on the NCBI. [1] The data was properly analyzed with the help of Microsoft Excel 2010.

As for the Excel opportunities, there were used 1) COUNTIFS function (for making a histogram), 2) MAX/MIN functions (for searching the biggest/shortest protein), 3) Pivot tables (for summarizing the number of genes on DNA strands).

3. Results



3.1 The analysis of lengths of proteins in the proteome of our bacterium.

During the work it was found out that most of encoded proteins have the length of 200-300 amino acid residues (aa). There are nearly 700 such proteins which comply with 24% of general number. The other huge amount of proteins – 650 – matches with the length of 100-200 aa. And also 500 proteins consist of 300-400 aa.

Starting with the length of 300 aa, we can see that as the length grows, the number of proteins falls considerably. So, the quantity of proteins with the length over 500 aa, reaches only ~ 14% of general amount.

The length of the biggest protein (*outer membrane adhesin-like protein*) is 3226 aa, while the shortest one (*hypothetical protein*) contains 32 aa.

3.2 The location of genes, encoding proteins(1) and RNA(2) on DNA strands

	(1)	(2)
Straight strand	1572	30
Complement strand	1392	19
Total	2964	49

Table 1. The distribution of genes of proteins and t/rRNA on straight and complement DNA strands.

On the straight strand there are located 1572 proteins which is 53% of total amount. On the other strand there are, 1392 proteins - 47% resp.

As for RNA-genes, the main part of them is distributed in the straight strand (61%) and other 39% are located on the complement one.

4. Discussion

The study has revealed that the biggest part of the proteome consists of proteins, which length changes between 200-400 amino acid residues. When the length gets over 400 aa. the number of proteins tends to decrease. The situation is quite typical for the prokaryotes.

The author express confidence that the genome of *Denitrovibrio acetiphilus DSM 12809* hides lots of mysterious facts that will be discovered in further researches.

5. Acknowledgements

The author wants to express gratitude to A.Alekseevskii and I.Rusinov for their great help in launching a study: for all materials, links and instructions given.

6. References

1. The national center for Biotechnology Information
<ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/>
2. [Supplementary materials](#)