Biochimica et Biophysica Acta 1844 (2014) 850–858

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Minireview

Molecular bases of protein halotolerance

Giuseppe Graziano [a], Antonello Merlino [b,c,*]

[a] Dipartimento di Scienze e Tecnologie, Università del Sannio, Benevento, Italy
[b] Dipartimento di Scienze Chimiche, Università di Napoli 'Federico II', Napoli, Italy
[c] Istituto di Biostrutture e Bioimmagini, CNR, Napoli, Italy

ARTICLE INFO

Article history:
Received 22 November 2013
Received in revised form 18 February 2014
Accepted 21 February 2014
Available online 1 March 2014

Keywords:
Halophile
Halotolerant protein
Molecular evolution
Protein stability
Protein structure
Salt adaptation

ABSTRACT

Halophilic proteins are stable and function at high salt concentration. Understanding how these molecules maintain their fold stable and avoid aggregation under harsh conditions is of great interest for biotechnological applications. This mini-review describes what is known about the molecular determinants of protein halotolerance. Comparisons between the sequences of halophilic/non-halophilic homologous protein pairs indicated that Asp and Glu are significantly more frequent, while Lys, Ile and Leu are less frequent in halophilic proteins. Homologous halophilic and non-halophilic proteins have similar overall structure, secondary structure content, and number of residues involved in the formation of H-bonds. On the other hand, on the halophilic protein surface, a decrease of nonpolar residues and an increase of charged residues are observed. Particularly, halophilic adaptation correlates with an increase of Asp and Glu, compensated by a decrease of basic residues, mainly Lys, on protein surface. A thermodynamic model, that provides a reliable explanation of the salt effect on the conformational stability of globular proteins, is presented.

1. Introduction

Understanding how the amino acid sequence of a protein determines its three-dimensional structure, dynamics, structural stability, and, ultimately, its biological function remains one of the most fundamental goals of structural biology. Although this problem is unlikely to be solved in a single step, one way of breaking it down is offered to us by nature's diversity itself. Adaptation to extreme environmental conditions, e.g., naturally extreme environments of temperature, pressure, pH, and salinity, led to the evolution of extremophilic enzymes that are able to be stable and to work in their respective hostile habitats, still retaining a significant degree of sequence identity and overall structural similarity to their mesophilic counterparts.

Examination of mesophilic/extremophilic enzyme pairs from different species side-by-side enables an assessment of how changes in sequence and structure have altered the enzyme properties [1]. Such comparisons also provide additional insights into structure–function relationship in general, and can be useful to give new impetus to the biotechnological research and to the development of improved enzymes for application in industrial processes. A wide variety of experimental and theoretical studies based on comparative investigation of thermophilic protein features with their mesophilic counterparts have been reported in literature (see for example [2–4]). Comparative structural analyses have also been performed to unveil the molecular bases of cold-adaptation [5–9]. Environmental adaptation of proteins at high salinity has been much less studied, although it is commonly known that there are numerous species that are inhabitants of extreme environments with excess salinity and that information about the survival mechanisms of halophilic proteins could well enable mesophilic enzymes to be modified to function efficiently in non-conventional solvents more relevant to the conditions used in many industrial processes [10–12].

This mini-review describes what is known about the molecular determinants of protein halotolerance; it focuses on the results of recent studies on structure and stability of halophilic and halotolerant enzymes (see below for the definitions). For the numerous biotechnological applications of halophilic and halotolerant proteins the reader is referred to other reviews [13–16].

2. Halophilic and halotolerant proteins

Halophiles are (micro)organisms that require salt for growth. They are adapted to live in high salinity environments (>0.2 M NaCl) and are resistant to the effect of osmotic stress. They include aerobic or anaerobic bacteria, cyanobacteria, archaea, protozoa, fungi, algae and multicellular eukaryotes. *Slight halophiles* grow optimally at 0.2–0.85 M (2–5%) NaCl, while *moderate halophiles* at 0.85–3.4 M (5–20%) NaCl

* Corresponding author at: Department of Chemical Sciences, University of Naples Federico II, Complesso Universitario di Monte Sant' Angelo, Via Cintia, I-80126 Napoli, Italy. Tel.: +39 081674276; fax: +39 081674090.
E-mail address: antonello.merlino@unina.it (A. Merlino).

http://dx.doi.org/10.1016/j.bbapap.2014.02.018
1570-9639/

and *extreme halophiles* at 3.4–5.1 M (20–30%) NaCl [17–19]. Our anthropocentric point of view led to the classification of halophilic (micro)organisms as extremophiles [20].

Proteins are called halophilic when they are isolated from halophiles. Halophilic proteins commonly need high NaCl concentrations to survive, halophilic enzymes require salt for catalytic activity.

There are also numerous proteins which tolerate high salt concentrations, but are not isolated from halophilic sources. These proteins are called halotolerant. Halotolerant enzymes remain active over a broad range of NaCl concentration, without any specific salt dependence [21,22].

Several independent studies have shown that halophilic and halotolerant proteins are able to work in conditions where ordinary enzymes may aggregate or lose enzymatic activity [21]. The peculiar ability of these proteins to be stable and to function in these extreme environments has been related to a number of different features. Main efforts to define the structural features of a protein that determine its resistance to this harsh environment are summarized below. In the following sections, sequence comparison at genome and proteome levels and structural comparison between halophilic proteins and their mesophilic counterparts will be described, with the aim to define the essential feature of a protein to cope with high salinity environment. Finally, a thermodynamic model for rationalizing halophilic protein stabilization at high salt concentration is discussed.

3. Structural adaptation: DNA level comparison

The first significant approach used to identify the structural determinants of protein halotolerance is based on comparative genome and proteome analysis of large data sets of sequences. A database dedicated to the sequenced halophilic genomes is available at HaloWeb: The Haloarchaeal Genomes Database website, http://halo4.umbi.umd.edu [23]. Other interesting information on sequenced halophilic genomes can be found at the "Halophile genome site", http://edwards.sdsu.edu/halophiles or at "HaloBase", http://halobase.info, a database for halophilic bacteria and archaea which covers molecular aspects and diversity based studies [24]. A representative list of sequenced halophilic genomes (August 2013) is reported in Table 1.

The genome of the first halophilic organism (*Halobacterium* sp. NRC-1) was sequenced in 2000 [25]. Analysis of its sequence highlighted a number of physical adaptations to high-salt environment: 1) proteome with an acidic pI (~4.5) [25,26]; 2) the lack of enzymes for the synthesis of at least eight amino acids; 3) the encoding of a group of signal transducers and transcriptional regulators. Comparative analyses with the sequence of a second halophile, *Haloarcula marismortui* (ATCC 43049), revealed general characteristics of halophilic genomes. In particular, shared features of halophilic genomes include an acidic proteome, multiple replicons (including a high G + C content), and multiple copies of general transcription factors [27].

A seminal work in this field is the comparative study of genome and proteome compositions of halophilic and non-halophilic microorganisms reported by Dutta and coworkers in 2008 [28]. According to these authors, at the genomic level, the dinucleotide abundance profiles of halophiles share some features which are rather distinctive from those of non-halophilic genomes. In particular, a high G + C content (well above 60%) has been observed, that presumably is useful to prevent UV-induced damages like the formation of thymine dimer and harmful mutations [23,29].

However, recently, it has been reported that not all halophiles bear this feature. For example, the overall G + C content of *Haloquadratum walsbyi* [30], *Halothermothrix oreni* [31] and *Salinicoccus kunmingensis* [32] is only between 42.9% and 47.9%. Thus, the G + C-bias does not seem to be a universal property of halophiles and other specific haloadaptation characteristics of nucleotide selection should be identified.

Table 1
Sequenced halophilic genomes.

Name	Strain or number	Genome size	Isolated from	Accession code
Halalkalicoccus jeotgali	DSM 18796	3.7 Mbp	From shrimp jeotgal, a traditional Korean fermented seafood	NCBI GenBank: CP002062.1–CP002068.1
Haloarcula hispanica	ATCC 33960	3.9 Mbp	From Spanish solar saltern	NCBI GenBank: CP002921, CP002922, CP002923
Haloarcula californiae	ATCC 33799	4.4 Mbp	Salt brine, Baja California, Mexico	Pubseed: 662475.4
Haloarcula marismortui	ATCC 43049	4.3 Mbp	From the Dead Sea	NCBI GenBank: AY59290–AY59298
Haloarcula sinaiiensis	ATCC 33800	4.4 Mbp	Salt brine, Israel	Pubseed: 662476.5
Haloarcula vallismortis	ATCC 29715	3.9 Mbp	Salt pools, Bad Water Point, Death Valley, CA	Pubseed: 662477.4
Halobacterium sp. NRC-1		2.6 Mbp	From salted food	NCBI GenBank: AE004437.1, AF016485.1, AE004438.1
Halobacterium salinarum R1		2.7 Mbp	Laboratory mutant, similar to *Halobacterium* sp. NRC-1	NCBI GenBank: AM774415.1–AM774419.1
Haloferax denitrificans	ATCC 35960		Saltern, California	Pubseed: 662478.4
Haloferax mediterranei	ATCC 33500	3.9 Mbp	Salt ponds, Alicante, Spain	Pubseed: 523841.6
Haloferax mucosum	ATCC BAA-1512	3.4 Mbp	Pastular mat from Hamelin pool, Shark Bay, Australia	Pubseed: 662479.5
Haloferax sulfurifontis	ATCC BAA-897	3.8 Mbp	Microbial mats and mineral crusts near the sulfide and sulfur rich Zodletone spring, southwestern Oklahoma, USA	Pubseed: 662480.4
Haloferax volcanii	ATCC 29605	4.0 Mbp	Shore mud, Dead Sea	NCBI GenBank: CP001953.1–CP001957.1
Halogeometricum borinquense	DSM 11551	3.9 Mbp	Solar saltern in Cabo Rojo, Puerto Rico	NCBI GenBank: CP001690.1–CP001695.1
Halomicrobium mukohataei	ATCC 700874	3.2 Mbp	Salt flat in Jujuy, Argentina	NCBI GenBank: CP001688.1–CP001689.1
Halophilic archaeon DL31		3.6 Mbp	From Deep Lake, Antarctica	NCBI GenBank: CP002988.1
Halopiger xanaduensis	SH-6	4.4 Mbp	From saline Lake Shangmatala in Inner Mongolia, China	NCBI GenBank: CP002839.1–CP002842.1
Haloquadratum walsbyi	DSM 16790	3.2 Mbp	From solar saltern	NCBI GenBank: AM180088.1–AM180089.1
Halorhabdus utahensis	DSM 12940	3.1 Mbp	Great Salt Lake in Utah, USA	NCBI GenBank: CP001687.1
Halorubrum lacusprofundi	ATCC 49239	3.7 Mbp	From Antarctic lake	NCBI GenBank: CP001365.1–CP001367.1
Haloterrigena turkmenica	ATCC 51198	5.4 Mbp	Sulfate saline soil in Turkmenistan	NCBI GenBank: CP001860.1–CP001866.1
Halovivax ruber XH-70		3.9 Mbp	From a saline lake in China	NCBI GenBank: CP003050.1
Natrialba magadii	ATCC 43099	4.4 Mbp	From Lake Magadi, Kenya	NCBI GenBank: CP001932.1–CP001935.1
Natronobacterium gregoryi SP2	ATCC 43098	3.8 Mbp	From Lake Magadi, Kenya	NCBI GenBank: CP003377.1
Natrononococcus occultus	ATCC 43101	4.0 Mbp	From Lake Magadi, Kenya	NCBI GenBank: CP003929.1
Natronomonas pharaonis	ATCC 35678	2.8 Mbp	From soda lake	NCBI GenBank: CR936257.1–CR936259.1

American Type Culture Collection (ATCC).