Biochimica et Biophysica Acta 1844 (2014) 497–504

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Molecular strategies to replace the structural metal site in the prokaryotic zinc finger domain

Ilaria Baglivo [a], Maddalena Palmieri [a], Alessia Rivellino [a], Fortuna Netti [a], Luigi Russo [a], Sabrina Esposito [a], Rosa Iacovino [a], Biancamaria Farina [b,c], Carla Isernia [a,b], Roberto Fattorusso [a,b], Paolo Vincenzo Pedone [a,b], Gaetano Malgieri [a,*]

[a] Department of Environmental, Biological and Pharmaceutical Sciences and Technologies, Via Vivaldi 43, 81100 Caserta, Italy
[b] Interuniversity Centre for Research on Bioactive Peptides (CIRPEB), University of Naples Federico II, Via Mezzocannone 16, 80134 Naples, Italy
[c] Institute of Biostructures and Bioimaging, CNR, Via Mezzocannone 16, 80134 Naples, Italy

ARTICLE INFO

Article history:
Received 7 August 2013
Received in revised form 20 December 2013
Accepted 24 December 2013
Available online 3 January 2014

Keywords:
Prokaryotic zinc finger
Thermal unfolding
DNA binding

ABSTRACT

The specific arrangement of secondary elements in a local motif often totally relies on the formation of coordination bonds between metal ions and protein ligands. This is typified by the ~30 amino acid eukaryotic zinc finger motif in which a β-sheet and an α-helix are clustered around a zinc ion by various combinations of four ligands. The prokaryotic zinc finger domain (found in the Ros protein from *Agrobacterium tumefaciens*) is different from the eukaryotic counterpart as it consists of 58 amino acids arranged in a ββββαα topology stabilized by a 15-residue hydrophobic core. Also, this domain tetrahedrally coordinates zinc and unfolds in the absence of the metal ion.

The characterization of proteins belonging to the Ros homologs family has however shown that the prokaryotic zinc finger domain can overcome the metal requirement to achieve the same fold and DNA-binding activity.

In the present work, two zinc-lacking Ros homologs (Ml_4 and Ml_5 proteins) have been thoroughly characterized using bioinformatics, biochemical and NMR techniques.

We show how in these proteins a network of hydrogen bonds and hydrophobic interactions surrogate the zinc coordination role in the achievement of the same functional fold.

1. Introduction

Structural zinc centers play a pivotal role in a large number of proteins contributing to a net stabilizing effect of the folded structure. The specific arrangement of secondary elements in a local motif often totally relies on the formation of coordination bonds between the metal ion and the protein ligands. This is the case of the ~30 amino acid zinc finger motif in which various combinations of four ligands recruit a zinc ion to stabilize the tertiary ββα fold of this domain; in the absence of the metal ion the domain unfolds.

The classical (or Cys_2His_2) zinc finger represents the most abundant DNA-binding domain in eukaryota [1–3] capable also of establishing protein–protein interactions [4,5]. Its structure is formed by a β-sheet and an α-helix clustered around the zinc ion by two cysteines and two histidine and by a three-residue hydrophobic core [1,6,7]. In prokaryotic organisms, the first Cys_2His_2 zinc finger was identified in the protein Ros from *Agrobacterium tumefaciens* [8]. Its interaction with DNA was studied in detail showing that the fundamental element for binding is a region comprising the Cys_2His_2 domain and four basic stretches, one positioned on the N-terminal side and three on the C-terminal side of the zinc finger domain [9].

The NMR structure of Ros DNA-binding domain (Ros87) has shown that the prokaryotic Cys_2His_2 domain tetrahedrally coordinates Zn(II) through the typical Cys_2His_2 coordination but exhibits a protein fold which is very different from that of the eukaryotic counterpart [10]. In particular, the Ros globular domain consists of 58 amino acids, arranged in a ββββαα topology and stabilized by an extensive 15-residue hydrophobic core [10]. More than 300 Ros homologues have been identified in a large number of bacteria, mostly belonging to the α subdivision of proteobacteria raising a number of interesting phylogenetic questions [11,12]. An inspection of Ros homologous sequences have shown that the zinc coordination residues are poorly conserved in Ros homologues even if their sequence identity can raise to 80% when matched to Ros87 sequence [13]. In particular, at least ten homologues of Ros are present in the bacterium *Mesorhizobium loti*. The second cysteine of the zinc coordination sphere of Ros is replaced by an aspartic residue in all the ten *M. loti* homologues, while the first histidine is always conserved. The first cysteine is replaced by a serine in three of the ten homologues and the fourth zinc-coordinating position of Ros is replaced by different amino acids [13]. The functional and structural properties of five of the ten *M. loti* homologues, representative of all the mutations found in the coordinating positions, have been investigated [13]. We have

* Corresponding author at: Via Vivaldi 43, 81100 Caserta, Italy.
E-mail address: gaetano.malgieri@unina2.it (G. Malgieri).

1570-9639/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.bbapap.2013.12.019

shown that, even if capable of specifically bind the same DNA sequence with a high affinity, two of them (named Ml_4 and Ml_5) bearing a serine in the first position did not contain a structural zinc ion (or any other metal ion). Therefore, the prokaryotic zinc finger domain in Ros homologues can either exploit a $CysAspHis_2$ coordination sphere, never described previously in DNA-binding zinc finger domains, or lose the metal while still retaining the same fold and DNA-binding activity. The presence of the zinc ion is strictly related to the presence of the cysteine in the first coordinating position [13]. The substitution of the serine in this position with a cysteine in one of the non-zinc-binding homologue makes the domain capable of coordinating zinc and does not profoundly affect the DNA-binding capability. Furthermore, the mutation of the first coordinating cysteine to serine converts a zinc-binding domain into a non-zinc-binding domain and, in the same way, does not affect DNA binding [13].

In order to elucidate at a molecular level how the prokaryotic zinc finger domain can overcome the metal requirement to achieve the same functional fold, the two proteins, Ml_4 and Ml_5, have been thoroughly characterized in the present work using bioinformatics, biochemical, and NMR techniques. We reveal how a network of hydrogen bonds and hydrophobic interactions surrogate the zinc coordination role in the achievement of the correct and functional fold.

These findings demonstrate a natural design strategy exemplifying how, during evolution [14], alternative strategies have been pursued within the same protein family to achieve the stabilization of a particular fold.

2. Materials and methods

2.1. EMSA

All of the proteins used for the DNA-binding experiments were expressed and purified from the *Escherichia coli* BL21 host strain as GST fusions, after cloning the coding sequence in the pGEX-6P-1 expression vector [13]. Purified proteins (7 pmol) were incubated with 2.5 pmol of the oligonucleotide 5′-GATTTTATATTTCAATTTTATTGTAA TATAATTTCAATTG-3′, in the presence of 25 mM HEPES (pH 7.9), 50 mM KCl, 6.25 mM $MgCl_2$, 1% NP-40, and 5% glycerol. After incubation, the mixture was loaded on a 5% polyacrylamide gel (29:1 acrylamide:bisacrylamide ratio) and run in 0.5× TBE at room temperature (200 V for 75 min). The gel was stained with Syber Green (Invitrogen), and the image was acquired using Typhoon Trio^{++} (Amersham).

2.2. Sample preparation

Uniformly, ^{15}N-labeled (^{15}N-Ml_4,^{15}N-Ml_4H37A, ^{15}N-Ml_5,^{15}N-Ml_5K27A,^{15}N-Ml_5H35A, ^{15}N-Ml_5Y40A and ^{15}N-Ml_5Y40F) proteins used for the NMR experiments were over-expressed and purified as previously published [9].

2.3. CD Experiments

Protein samples were prepared in 20 mM phosphate buffer containing 0.2 M NaCl adjusted at pH 6.8. The thermal denaturation of the proteins (Table 2) was evaluated using a JASCO J-815 CD spectropolarimeter equipped with Peltier temperature control. CD spectra were measured at regular temperature intervals in the 277–383 K range. After the final measurement, the samples were cooled to 277 K, and a final set of spectra was collected. The samples were generally 25 μM, and data were collected using a quartz cuvette with a 1-cm path length in the 200- to 260-nm wavelength range with a data pitch of 1 nm. All data were recorded with a bandwidth of 1 nm with a scanning speed of 50 nm/min and normalized against reference spectra to remove the background contribution of buffer. The data obtained were fitted into two-state folding model [15].

2.4. NMR experiments

NMR samples contained 1 mM of uniformly ^{15}N-labeled proteins, 20 mM phosphate buffer adjusted at pH 6.8, 0.2 M NaCl, and 90% H_2O/10% ^2H_2O. All the NMR experiments were acquired on a Varian Unity INOVA 500-MHz spectrometer. NMR experiments were processed by using Varian (VNMR 6.1B) software. ^{1}H and ^{15}N chemical shifts were calibrated indirectly by using TMS as external references. The program XEASY [16] was used to analyze the spectra. The structures were analyzed and visualized by using the program MOLMOL [17]. For the thermal unfolding experiments, a series of ^{1}H-^{15}N HSQC spectra were acquired increasing temperatures at regular intervals of generally 3 K from 277 to 353 K with the following parameters: the number of complex points were 256 for ^{15}N (F1) and 2048 for ^{1}H (F2). Ml_4 and Ml_5 have only one tryptophan each, which is a key component of the hydrophobic core and crucial for these proteins folding. The intensity variation of indolic NH signal of the Trp *versus* temperature was monitored and then fitted using a non-linear least-squares method to obtain unfolding curves reported and T_m [18]. All of the unfolding curves were obtained using the program GraphPad Prism5.

3. Results

This work will focus the attention on the proteins $Ml4_{56–151}$ and $Ml5_{56–154}$, N-terminal deletion mutants of Ml_4 and Ml_5 corresponding to Ros87 DNA-binding domain (Ros87) [9,13,19]. They are zinc-lacking proteins able to bind the same Ros DNA target sequence with high affinity as demonstrated by ICP-MS data, band shift, and NMR experiments [13]. Moreover, as this article deals also with mutants of these proteins, from now on, the name of $Ml4_{56–151}$ and $Ml5_{56–154}$ will be shortened to Ml_4 and Ml_5 and the numbering of their amino acids will start from 1 to simplify the task to readers who are not familiar with them. An alignment of the two proteins to the Ros87 primary sequence is reported in Fig. 1. The residues found in those that are zinc-coordinating positions in Ros87 are S_{24}, D_{27}, H_{37}, and G_{42} in Ml_4 and S_{22}, D_{25}, H_{35}, and Y_{40} in Ml_5 (Fig. 1).

3.1. Bioinformatics analysis: the structure of Ml_4 and Ml_5

A high reliable structural model of Ml_4 using the CS23D server [20] has been obtained exploiting the complete ^{1}H, ^{15}N, and ^{13}C NMR chemical shift assignment of Ml_4 previously reported by us [21].

Ml_4 tertiary fold shows a globular domain with a ββββαα topology and, though lacking the zinc ion center, its overall structure is similar to that of Ros87. This was anticipated by their high sequence identity (more than 60% within the folded region), the similar chemical shifts of the conserved residues, and their similar DNA-binding activity [13,21,22].

The 3D structural model of Ml_4 was here also obtained with *i-TASSER* [23] using the primary sequence as input data. *i-TASSER* models are built based on multiple-threading alignments and iterative template fragment assembly simulations.

The two models (obtained with *CS23D* and *i-TASSER*) show a good conservation of the secondary structure, with an RMSD for the globular folded region of 1.33 Å (Fig. 2). They differ slightly in the first α-helix and the turn between the first and the second α-helices but the overall fold, distribution, and orientation of secondary structural elements, as well as the arrangement of the side chains of the residues in the zinc-coordinating equivalent sites, are conserved. For these reasons, as the protein has not been previously characterized via NMR, the *i-TASSER* server was used to obtain Ml_5 structural model (Fig. 3).

As expected, Ml_5 shows a globular domain with a ββββαα topology and an overall structure similar to that of Ros87, just as Ml_4. Predictable minor differences are found between the Ml_4 and the Ml_5 structures, being local rearrangements the response to mutations to preserve the global fold [24].