PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Solution and crystal structure of BA42, a protein from the Antarctic bacterium *Bizionia argentinensis* comprised of a stand-alone TPM domain

Martin Aran,[1] Clara Smal,[1] Leonardo Pellizza,[1] Mariana Gallo,[1] Lisandro H. Otero,[1,2] Sebastián Klinke,[1,2] Fernando A. Goldbaum,[1,2] Esteban R. Ithurralde,[3] Andrés Bercovich,[4] Walter P. Mac Cormack,[5] Adrián G. Turjanski,[3] and Daniel O. Cicero[6]*

[1] Fundación Instituto Leloir, IIBBA-CONICET, Patricias Argentinas 435 (C1405BWE), Buenos Aires, Argentina

[2] Plataforma Argentina de Biología Estructural y Metabolómica PLABEM, Patricias Argentinas 435 (C1405BWE), Buenos Aires, Argentina

[3] Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, e INQUIMAE-CONICET, Intendente Güiraldes 2160 (C1428EGA), Buenos Aires, Argentina

[4] Biosidus S.A., Constitución 4234 (C1254ABX), Buenos Aires, Argentina

[5] Instituto Antártico Argentino, Cerrito 1248 (C1010AAZ), Buenos Aires, Argentina

[6] Dipartimento di Scienze e Tecnologie Chimiche, Università di Roma "Tor Vergata", via della Ricerca Scientifica SNC (00133), Rome, Italy

ABSTRACT

The structure of the BA42 protein belonging to the Antarctic flavobacterium *Bizionia argentinensis* was determined by nuclear magnetic resonance and X-ray crystallography. This is the first structure of a member of the PF04536 family comprised of a stand-alone TPM domain. The structure reveals a new topological variant of the four β-strands constituting the central β-sheet of the αβα architecture and a double metal binding site stabilizing a pair of crossing loops, not observed in previous structures of proteins belonging to this family. BA42 shows differences in structure and dynamics in the presence or absence of bound metals. The affinity for divalent metal ions is close to that observed in proteins that modulate their activity as a function of metal concentration, anticipating a possible role for BA42.

Proteins 2014; 82:3062–3078.
© 2014 Wiley Periodicals, Inc.

Key words: structural genomics; Antarctic bacteria; *Bizionia argentinensis*; BA42; protein structure; nuclear magnetic resonance; X-ray crystallography.

INTRODUCTION

The family *Flavobacteriaceae* comprises more than 40 genera, many of them constituted by strains recovered from a variety of marine habitats. Particularly interesting are those living in polar marine environments, where they play a relevant role in the mineralization of organic matter.[1] Recently, an orange-pigmented psychrophilic marine bacterial strain designated JUB59 was isolated from Antarctic surface seawater. Further biochemical and phylogenetic characterization proved that strain JUB59 represents a novel species of the genus *Bizionia* into the family *Flavobacteriaceae* for which the name *Bizionia argentinensis* sp. nov. was proposed.[2]

The genome of *B. argentinensis* was recently decoded[3] and has been deposited at the DDBJ/EMBL/GenBank under the accession code AFXZ00000000. This genome constitutes a relevant source for the discovery of new proteins showing biological activity in extreme conditions of low temperatures. These biomolecules can significantly contribute to the understanding of the general mechanisms allowing biological systems to adapt themselves to these particular life conditions.

Additional Supporting Information may be found in the online version of this article

Grant sponsor: Argentinian Research Council (CONICET); Grant sponsor: Agencia Nacional de Promoción de la Ciencia y la Tecnología; Grant number: PID-2010 N°0006.

*Correspondence to: Daniel Oscar Cicero, via della Ricerca Scientifica SNC (00133), Rome, Italy. E-mail: cicero@scienze.uniroma2.it

Received 13 June 2014; Revised 1 August 2014; Accepted 6 August 2014
Published online 13 August 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24667

 © 2014 WILEY PERIODICALS, INC.

Seqencing of this new genome also provides an opportunity to investigate the function of genes with unknown function that are largely conserved within the family *Flavobacteriaceae*. The emergence of novel DNA sequencing technologies in recent years has resulted in an enormous, and ever-growing, number of gene sequences.[4] Despite the large increase in the number of sequences, the overall structural coverage of proteins—including proteins for which reliable homology models can be generated—on the residue level has increased from 30% to 40% over the last 10 years.[5] Three-dimensional structural information often provides the insight required to understand macromolecular function and to progress toward a greater understanding of proteomes through analysis of macromolecular complexes and assemblies.[6–8] Comprehensive structural coverage and accurate functional annotation are universal goals that are being actively pursued by a wide range of individual investigators, as well as by multidisciplinary, multiinstitutional consortia.[5] Without discovery-driven efforts, large amounts of biology will continue to remain obscure, and it will not be possible to systemically examine the structural features of full proteomes, which undoubtedly would lead to unexpected functional and biological insights.

A protocol based on bioinformatics analysis and NMR screening was set up to identify suitable targets for structure determination among the proteins coded in the *B. argentinensis* genome presenting features that make them of interest in a structural genomics project: cytosolic location, low sequence identity with proteins of known function, presence of homologs in the family *Flavobacteriaceae*, and low to medium molecular weight. Based on these criteria, 50 ORFs from the *B. argentinensis* genome were selected, subsequently amplified by PCR and cloned in expression vectors. Further analysis of protein expression and solubility in *Escherichia coli* BL21 cells allowed the identification of more than 30 candidates for structure determination. The first selected target was BA42, a 145-residue long protein coded by the ORF 42 contig 3, which showed excellent solubility and expression levels in *E. coli*.[9]

A BLAST search revealed that BA42 gives the highest hits (E-value $\leq 1 \times 10^{-46}$) to flavobacterial proteins only (Fig. 1). The closest sequence homolog shows only a 67% amino acid identity, indicating that protein sequences of organisms that are closer in the evolutionary line might remain undiscovered. BA42 encodes a single domain constituted by a TPM (TLP18.3/Psb32/MOLO-1; PFAM PF04536, formerly known as DUF477), a protein family with more than 1350 members in prokaryotic and eukaryotic kingdoms. Very recently, the NMR structure of two TPM domains belonging to prokaryotic proteins (CG2496 from *Corynebacterium glutamicum* [PDB code 2KPT] and PG0361 from *Porphyromonas gingivalis* [PDB code 2KW7])[10] and the X-ray structure of the TPM domain of *Arabidopsis thaliana* AtTPL18.3 protein (PDB code 3PTJ)[11] have been determined. These three structures, despite sharing a very low sequence identity, present a similar αβα "sandwich" structure, not previously found in other protein domains. The function of CG2496 and PG0361 remains unknown, whereas AtTLP18.3 was classified as a thylakoid acid phosphatase,[11] and was proposed to provide partial protection for the PSII complex to resist photoinhibition under fluctuating high-light conditions.[12] In this context, we have solved the structure in solution and in the crystalline state of BA42 to gain information about its possible functional role. Our results show that BA42 presents a circular permutation of the common fold observed for the other three members of the PF04536 family, and two metal binding sites for divalent metal ions, which stabilize a rare structural feature of two crossing loops.

MATERIALS AND METHODS

Cloning

The BA42 gene was amplified by PCR using *B. argentinensis* genomic DNA as template, 5′-GAGAACCTGTA CTTTCAGGGTATGTCTAAA ATAGAAGAGTTTTTAACA G-3′ as 5′-primer and 5′-GGGGACCACTTTGTACAAGA AAGCTGGGTTATTATCCTTTTGAAATTG-TGTTTGGTA ATTC-3′ as 3′-primer. The PCR product was purified, amplified with the forward primer 5′-GGGGACAAGTTT GTACAAAAAAGCAGGCTCGGAGAA CCTGTACTTTCA G-3′ and the reverse primer 5′-GGGGACCACTTTGTAC AAGAAAGCTGGGTTA-3′, and recombined using the Gateway® BP Clonase® II enzyme mix into the pDONR vector (Invitrogen). The plasmid DNA from positive clones was purified, confirmed by DNA sequencing, and recombined into the pDEST-527 expression vector using the Gateway® LR Clonase® II enzyme mix.

Protein expression and purification

For ^{15}N- and ^{13}C-labeling, BL21 (λDE3) *E. coli* cells harboring the BA42 expression plasmid were grown at 37°C in M9 medium containing 1 g L^{-1} $^{15}NH_4Cl$, 2 g L^{-1} U-$^{13}C_6$-glucose, and 100 μg mL^{-1} ampicillin up to an $OD_{600\ nm} \sim 0.8$. Recombinant protein expression was induced with 1 m*M* isopropyl-β-D-thiogalactopyranoside for 16 h at 20°C. Cells were centrifuged (5000*g*, 10 min) and the pellet resuspended in 20 m*M* Tris–HCl (pH 7.5), 0.5*M* NaCl, 40 μg mL^{-1} phenylmethylsulfonyl fluoride and disrupted by sonication. The cell debris was clarified by centrifugation (40,000*g*, 30 min). The soluble fraction containing the *6xHis*-tagged protein was loaded onto a HiTrap IMAC HP column (GE Healthcare). The column was washed with 10 volumes of binding buffer (20 m*M* Tris–HCl, 0.5*M* NaCl, pH 7.5) and the recombinant protein was eluted using a linear gradient of up to 100%

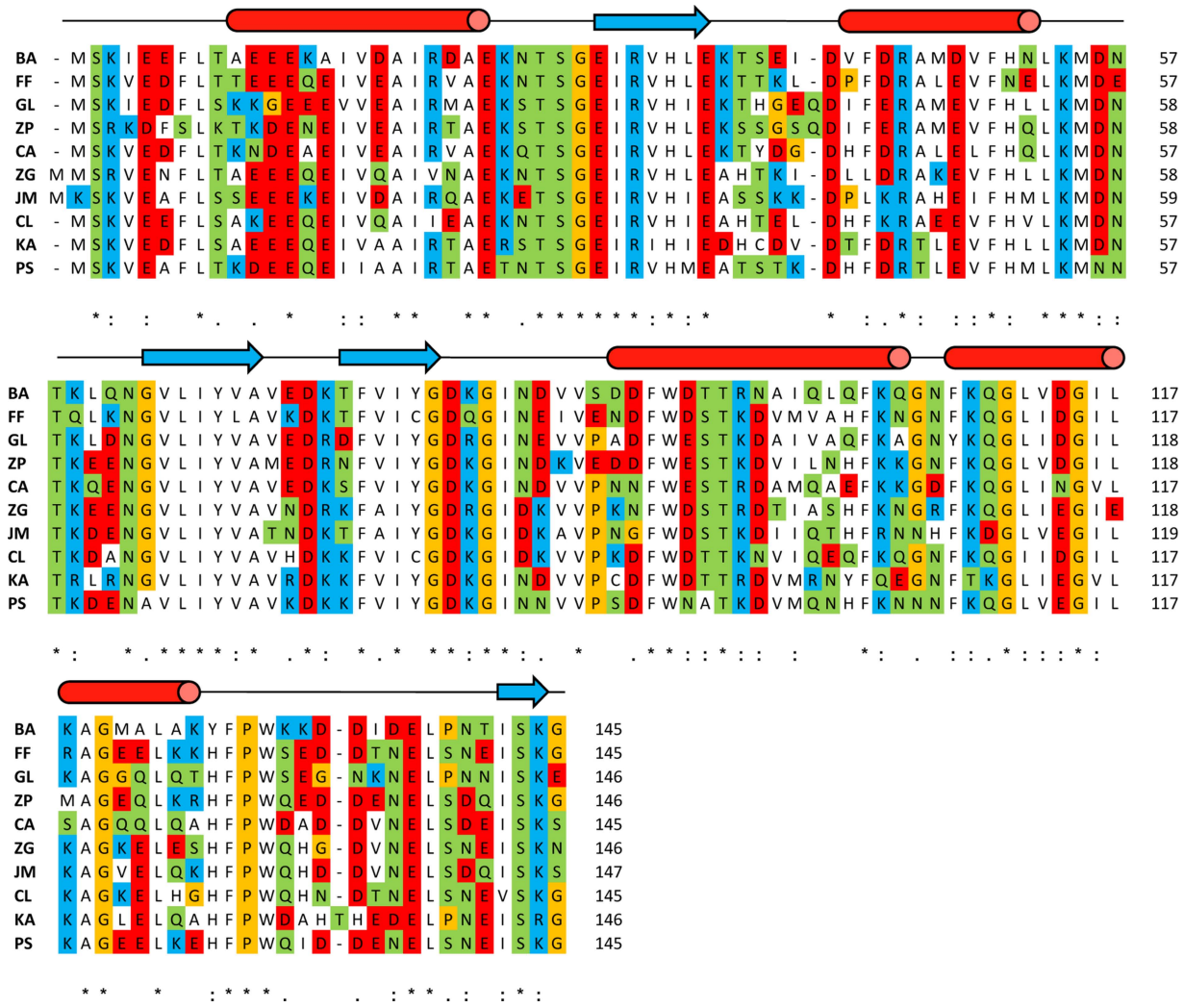

Figure 1

Multiple alignment of BA42 with amino acid sequences of close sequence homologs generated by ClustalW (http://www.ebi.ac.uk/clustalW). Amino acids are shaded by property: blue, positively charged; red, negatively charged; white, hydrophobic and aromatic; green, neutral hydrophilic; orange, glycine or proline. The species abbreviations and NCBI reference sequence codes are: BA: *Bizionia argentinensis* (ZP_08818491.1), FF: *Flavobacterium frigoris* (ZP_09895063.1), GL: *Gillisia limnaea* (ZP_09670103.1), ZP: *Zunongwangia profunda* (YP_003583329.1), CA: *Croceibacter atlanticus* (YP_003715719.1), ZG: *Zobellia galactanivorans* (YP_004735430.1), JM: *Joostella marina* (ZP_10108153.1), CL: *Cellulophaga algicola* (YP_004163638.1), KA: *Kordia algicida* (ZP_02161151.1), PS: *Polaribacter* sp. (ZP_01053120.1). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

elution buffer (20 m*M* Tris–HCl, 0.5*M* NaCl, 0.7*M* imidazole, pH 7.5). The protein fraction collected in the elution step was dialyzed against 20 m*M* Tris–HCl, 0.1*M* NaCl (pH 7.5) and the His-tag was cleaved using TEV protease (Sigma–Aldrich) with 0.03 mg mg^{-1} fusion protein at 28°C for 16 h. After cleavage, the TEV protease and His-tag were removed by passage through a HiTrap IMAC HP column. The purified recombinant BA42 protein was concentrated and subjected to gel filtration using a Superdex-75 column (GE Healthcare). The purity of the eluted protein sample was assayed by SDS-PAGE. The absorbance at 280 nm was used to estimate the protein content of homogeneous BA42 preparations.

NMR measurements

NMR spectra were collected at 301 K on a Bruker Avance 600 MHz spectrometer equipped with a triple-resonance cryoprobe. Chemical shifts for BA42 were already available.[9] For the structure determination, three-dimensional $^{1}H—^{15}N$ NOESY and $^{1}H—^{13}C$ NOESY, separately optimized for aliphatic and aromatic side chains, were collected with 100 ms mixing time. Three distinct sets of residual dipolar constants (RDCs) were measured: $^{1}J_{HN-N}$, $^{3}J_{H\alpha-C\alpha}$, $^{1}J_{C\alpha-C'}$. Anisotropic data were collected by dissolving the protein directly in 14 mg mL^{-1} liquid crystalline Pf1 (Asla Biotech) as the alignment medium.

$^1J_{H\alpha-C\alpha}$ couplings were obtained from the ^{1}H-coupled (F1) version of the HACACO experiment.[13] $^1J_{HN-N}$ RDCs were measured by an in-phase/anti-phase HSQC experiment.[14] $^1J_{C\alpha-C'}$ RDCs were measured in a modified version of the *J*-modulated HN(COCA) experiment.[15] NMR data were processed with NMRPipe[16] and analyzed with NMRView[17] programs on a Linux workstation. Spectra were zero filled to 4K and 0.5K points in the *t*2 and *t*1 dimensions, respectively, and were apodized using a phase-shifted Gaussian function.

The ^{1}H and ^{15}N resonance variations were followed at 35°C by collecting HSQC experiments. Combined chemical shift perturbations were calculated with a scaling factor (α_N) of 0.17 using the following equation:

$$\Delta \text{ ppm} = \text{sqrt}\ ((\Delta\delta_{HN})^2 + (\Delta\delta_N \times \alpha_N)^2).$$

Hydrodynamic measurements of BA42 using pulsed-field gradients

Diffusion-ordered spectroscopy experiments were performed using the BPPSTE (bipolar pulse pairs stimulated echo) method.[18] The duration of the total diffusion phase-encoding gradient pulse was 2 ms, the diffusion delay was 0.15 s, and the minimum gradient strength was set from 0.25 to 50 G cm^{-1}. Diffusion coefficients were measured using 0.4 m*M* BA42 in 10 m*M* phosphate buffer (pH 7.5) at 28°C, and 0.2% dioxane was used as a radius standard.[19]

^{15}N relaxation and ^{1}H—^{15}N heteronuclear NOE measurements of BA42

The ^{15}N relaxation experiments were all measured at 28°C. Measurements of ^{15}N T_1 and ^{15}N T_2 were performed at 60.81 MHz ^{15}N frequency (1.5 m*M* protein concentrations at pH 7.5) using standard pulse schemes.[20,21] Six time points were collected for ^{15}N T_1 and ^{15}N T_2 in an interleaved manner. The ^{15}N T_1 delays used were of 14, 350, 700, 1050, 1400, and 2100 ms, and of 8.4, 25.1, 41.9, 75.4, 125.7, and 167.6 ms for ^{15}N T_2. In the direct dimension (^{1}H), the carrier frequency was set on the water resonance with a spectral width of 22.7 ppm; the indirect (^{15}N) dimension was centred at 116.8 ppm with a spectral width of 25 ppm. Two thousand complex data points with 256 complex increments were collected. Typical T_1 experiments were acquired with 64 scans and a repetition delay of 3 T_1s, while T_2 experiments were acquired with 64 scans.

^{1}H—^{15}N heteronuclear NOE data were acquired with interleaved, unenhanced, and enhanced (3 s ^{1}H excitation period) spectra. ^{1}H—^{15}N heteronuclear NOE values represent the ratio of peak heights from the enhanced and unenhanced spectra.

NMR structure calculation

Structure calculation was performed with XPLOR-NIH[22] using a simulated annealing protocol,[23] slightly modified to incorporate RDCs, and a refinement against a database of Ramachandran plot dihedral angles.[24,25] Calculations were performed over residues 7–145, excluding the first six N-terminal residues that turned out to be unstructured. The simulated annealing was driven by NMR-derived constraints during the entire calculation, with force constraints and weighting factors optimized to increase the score of low energy structures. A set of 2313 interproton distance restraints was used, subdivided into three groups: strong (1.8–3.5 Å), medium (1.8–4.5 Å), and weak (3.2–5.0 Å). Backbone dihedral restraints (φ and ψ angles) were derived from chemical shift analyses using the program TALOS+.[26] Backbone hydrogen bonds were recognized by evaluating the spatial relationship of amide protons with potential acceptors in the initial structures produced without the use of hydrogen bond constraints. In addition, specific side chain hydrogen bonds were clearly identified using the WHAT IF server[27] with structures generated in later stages of the structural determination and were used as additional restraints in the last cycle of minimization. Three distinct sets of backbone RDCs (Table I) were included. Observed RDCs were fitted to a preliminary structure, using the singular value decomposition method with the PALES program,[28] which provides an initial guess of the magnitude and orientation of the molecular alignment tensor. Subsequent optimization during the simulated annealing process led to the following final values for the N—H normalized magnitude of the RDC tensor and rhombicity, respectively: DaNH = 5.6 Hz and R = 0.66. A complete cross-validation of RDCs was conducted by carrying out a series of simulated annealing calculations, each one lacking 10% of RDCs randomly chosen from the whole dataset. The missing RDCs were back-calculated to evaluate how well each RDC could be predicted; an average *R*free of 32.75% was obtained. The 20 lower energy target function structures (from a total of 1000) showing no distance violations above 0.5 Å and dihedral violations above 5° were subjected to an energy minimization process. The energy minimization procedure has been performed through the Gromacs 4.6.5 MD package.[29] The 20 initial system topologies have been obtained through the pdb2gmx program and parameterized with the AMBER ff03 force-field.[30] The structures have been immersed in a triclinic box filled with TIP3P water molecules[31] imposing a minimum distance of 1.0 nm between the solute and the box walls, while the charges have been neutralized adding 9 Na^+ counterions. A total of 2000 minimization steps, using the conjugate gradient algorithm, have been executed for each of the 20 systems, imposing restraints of 1000 kJ mol^{-1}nm^{-1}

Table I
Experimental Restraints and Structural Statistics for the 20 Lowest-Energy Structures of the BA42 Protein

Number of experimental restraints	2650
Distance restraints from NOEs	1560
Intraresidue	437
Sequential	498
Medium range	273
Long range	352
Hydrogen bond distance restraints	172
Dihedral angle restraints	254
Residual dipolar coupling constants	332
H—N	116
H_α—C_α	113
C_α—CO	103
Average number of restraints per residue	19
XPLOR energies (kcal mol^{-1})	
E_{total}	613.2 ± 19.8
E_{bond}	23.0 ± 3.7
E_{angle}	282.2 ± 7.8
$E_{improper}$	58.1 ± 3.0
E_{vdw}	26.9 ± 8.1
E_{cdih}	0.3 ± 0.2
E_{noe}	25.6 ± 2.7
E_{sani}	197.1 ± 3.0
Rms deviations from experimental restraints	
Average distance restraints violation (Å)	0.021 ± 0.001
Average dihedral angle restraints violation (°)	0.13 ± 0.06
Average H—N RDC constants violation (Hz)	0.68 ± 0.03[a]
Average H_α—C_α RDC constants violation (Hz)	1.06 ± 0.03[a]
Average C_α—CO RDC constants violation (Hz)	3.13 ± 0.05[a]
R_{free} (%)	
H—N RDC	30.8
H_α—C_α RDC	26.8
C_α-CORDC	25.9
Rms deviations from idealized covalent geometry	
Bond (Å)	0.0030 ± 0.0001
Angle (°)	0.660 ± 0.007
Improper (°)	0.57 ± 0.01
MolProbity analysis	
Clashscore[b]	0
Ramachandran favored (%)	95.8 ± 1.2
Ramachandran outliers (%)	0.39 ± 0.48
MolProbity score[c]	0.99 ± 0.22
What if quality indicators[d,e]	
First generation packing quality	1.9 ± 0.5
Second generation packing quality	5.1 ± 1.5
Ramachandran plot appearance	1.7± 0.2
Chi-1/chi-2 rotamer normality	3.5 ± 0.5
Backbone conformation	0.5 ± 0.2
Coordinates precision	
	Backbone (Å) / All heavy atoms (Å)
	0.27 ± 0.04 / 0.84 ± 0.06

[a]Normalized with respect to the H—N value.
[b]Number of serious steric overlaps (>0.4 Å) per 1000 atoms.
[c]Combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.
[d]Structure *Z*-scores, positive is better than average.
[e]Residues 7–145.

on the protein backbone atoms and of 500 kJ mol^{-1} nm^{-1} on the sidechains atoms, while the water atoms have been considered as fully flexible, to remove all the unfavorable interaction without altering the solved structures. No experimental restraints were used in this final minimization. An ensemble of the 20 minimized models was chosen to represent the protein solution structure and, together with the constraint list, they were deposited in the Protein Data Bank (code 2MPB). Structural statistics are provided in Table I.

Crystallization

Initial crystallization trials were made on 96-well plates using a Honeybee 963 robot (Digilab, Marlborough, MA) and screens from Jena Bioscience (Jena, Germany) and Hampton Research (Aliso Viejo, CA), in a sitting drop vapor diffusion configuration. After 5 days of equilibration at 20°C, 22 conditions out of the 480 tested showed promising hits, most of them corresponding to tiny needles. Crystals were optimized by the hanging drop method by mixing 1 μL of the concentrated protein stock (16 mg mL^{-1} in Tris 10 m*M*, pH 7.5) with an equal amount of crystallization solution. The following condition yielded the best crystals: 24% (w/v) PEG 2000 MME, 0.1*M* sodium acetate, 0.1*M* MES, pH 6.5. In this experiment, long thin bars with a maximum size of 0.4 mm × 0.05 mm × 0.05 mm were obtained after 1 week. Samples were cryoprotected in mother liquor added with 10% (w/v) PEG 400 and then cryocooled in liquid nitrogen in Hampton Research loops.

X-ray data collection and processing

Diffraction data were collected at 100 K on a single crystal at the PROXIMA 1 protein crystallography beamline at the SOLEIL Synchrotron (France), using a PILATUS 6M detector (Dectris, Baden, Switzerland). Data were processed with the program XDS[32] running in a semi-automated approach using the xdsme package (http://code.google.com/p/xdsme/). A total of 5% of the recorded reflections were flagged for cross validation. Details on data collection parameters and processing statistics are shown in Table II.

Structure resolution, refinement, and deposition

The BA42 structure was solved by the molecular replacement method with the program AMoRe,[33] using the coordinates of the previously solved NMR model as starting probe. A unique straightforward solution was found in the asymmetric unit, and the oriented coordinates were subjected to positional and individual B-factor refinement with the program Buster.[34] Hydrogen atoms were included with the MolProbity server,[35] and model building was performed with the program Coot.[36] In the last refinement cycles, 169 water molecules as well as two Ca^{2+} ions were added to the crystallographic model. The final coordinates were deposited in the Protein Data Bank under the code 4OA3. Details on the refinement process are found in Table II.

Table II
Experimental Data and Structural Statistics of the Crystal Structure of BA42

		BA42
Data collection		
Synchrotron source		SOLEIL
Beamline		PROXIMA 1
Number of frames		600
Oscillation step (°)		0.2
Detector distance (mm)		196
Wavelength (Å)		0.97857
Exposure per frame (s)		0.2
Indexing and scaling		
Cell parameters	*a* (Å)	31.92
	b (Å)	39.62
	c (Å)	106.97
	$\alpha = \beta = \gamma$ (°)	90
Spacegroup		$P2_12_12_1$
Resolution limit (Å)		1.39
Number of total reflections		118,799
Number of unique reflections		27,591
Average multiplicity [a]		4.3 (4.4)
$<I/\sigma(I)>$		16.0 (2.0)
R_{merge}		0.049 (0.695)
Completeness (%)		98.1 (96.8)
Molecules per asymmetric unit		1
Solvent content (%)		40
B-factor (Wilson plot, Å^2)		20
Refinement		
Resolution range (Å)		30.6–1.39
Number of protein atoms		1109
Number of ligand atoms		2
Number of water molecules		169
R		0.174
R_{free}		0.205
Rms deviations from ideal values		
Bond lengths (Å)		0.01
Bond angles (°)		1.1
B-factor (average, Å^2)		24
MolProbity validation[b]		
Clashscore		0.88
Poor rotamers (%)		1.6
Ramachandran plot		
Favored (%)		98.6
Allowed (%)		1.4
Disallowed (%)		—

[a]Values in parentheses correspond to the highest resolution shell: 1.47–1.39 Å.
[b]Chen *et al.*[35].

Circular dichroism experiments

Far-UV circular dichroism (CD) measurements were performed on a Jasco J-810 spectropolarimeter using a Peltier temperature-controlled sample holder in a 0.2 cm path length cell with a protein concentration of 5 μ*M* in presence of 20 m*M* Tris–HCl pH 7.0. Thermal unfolding curves were followed by monitoring at 223 nm using 2.0°C min^{-1} ramp rate, 1.0°C of sampling with 5 s of waiting. The metal-free sample was measured in addition of 1 m*M* of EDTA and the metal-bound sample was performed by incubating with 30 μ*M* $CaCl_2$.

RESULTS

BA42 shows a three-layered αβα architecture in solution

BA42 was obtained as a recombinant protein and labeled with $^{15}N/^{13}C$ as previously described.[9] In this work, we tackled the determination of the BA42 structure. Our first goal was to determine the multimerization state of BA42 in solution. BA42 behaves as a globular monomeric protein as judged from the experimental hydrodynamic radius (Rh = 20.3 ± 0.2 Å), obtained using a PFG experiment.[37] The observed value fits well with that expected for a well-folded protein of 145 residues (20.1 Å).[38]

The high-resolution structure of BA42 in solution was determined by NMR standard procedures, including the use of three types of residual dipolar couplings (RDCs). BA42 shows a defined conformation for residues 8–145. Residues 4–7 showed a lower T_1/T_2 ratio than the average, indicating the existence of a local flexibility in the ps time-scale (see below). A good convergence was achieved for the 8–145 region, resulting in a 0.40 ± 0.05 Å and 1.12 ± 0.06 Å rmsd for backbone and heavy atoms with respect to the average, respectively [Fig. 2(A)]. Quality indicators and the distribution in the Ramachandran regions are in line with those of current reliable structures (Table I).

The structure of BA42 shows an α/β fold [Fig. 2(B)] consisting of five α-helices and a four-stranded β-sheet with the topology D(↑)A(↑)B(↑)C(↓). α-Helices I, IV, and V are packed against one side of the β-sheet, while α-helix II is located on the opposite side, constituting an αβα sandwich structure. α-Helix III is perpendicular to the plane of the middle layer. The locations of regular secondary structure elements are: β-strands A (residues 30–36), B (64–70), C (75–80), and D (142–144), and α-helices I (10–23), II (43–53), III (82–85), IV (94–104), and V (109–127).

BA42 shows a rigid conformation in solution

^{15}N relaxation rates measurements were used to determine the global and local dynamics of BA42. The correlation time of BA42 (145 aa) calculated using the T_1/T_2 average value is 8.0 ± 0.2 ns at 28°C, in accordance with the prediction of a simple empirical formula for a globular protein of this MW in water at 300 K.[39] This result, together with the observed value of the hydrodynamic radius already discussed, indicates that BA42 behaves as a monomer in solution, even at the high concentrations used for the structural study (around 0.8 m*M*).

Figure 3(A) shows the experimental ^{15}N T_1/T_2 ratios as well as those calculated for the lowest energy structure considered as a rigid body, using the program HYDRONMR.[40] Comparison of both values in the

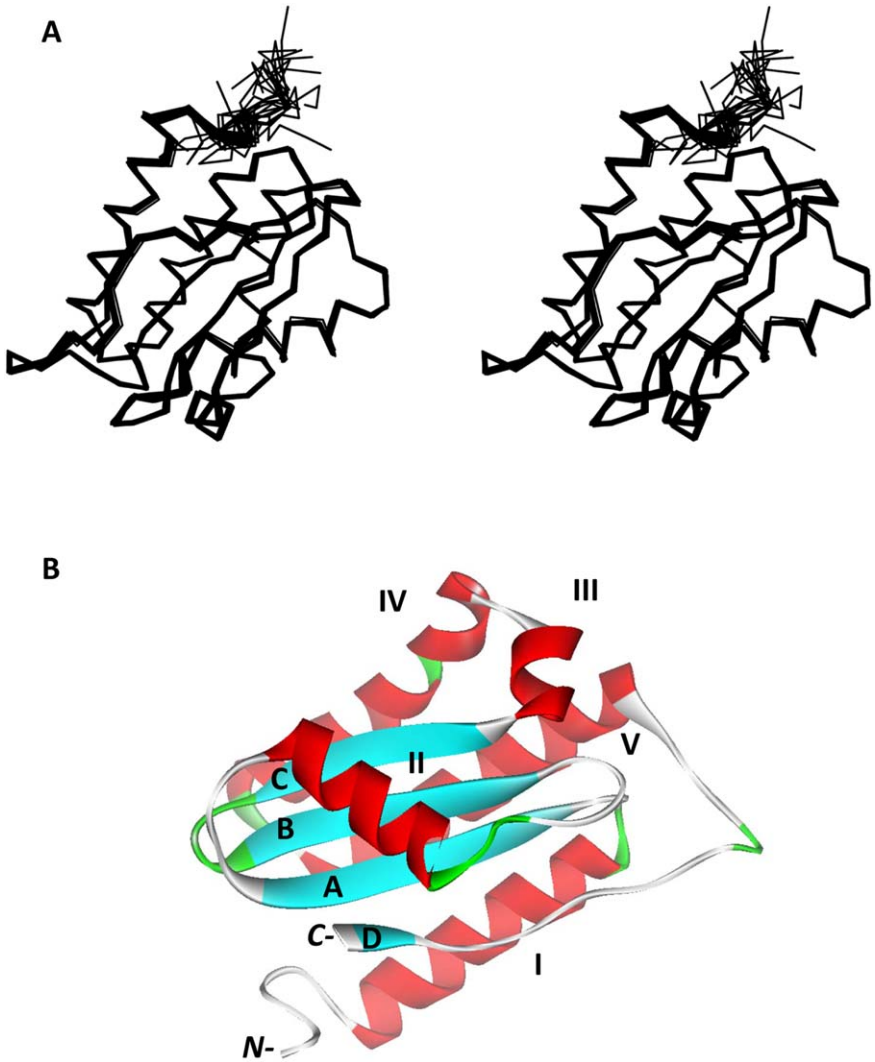

Figure 2

Solution structure of BA42. (**A**) Stereoview of the solution structure of BA42. The bundle of 20 energy minimized structures with the lowest target function is shown. (**B**) Ribbon view of the structure closest to the mean, α-helices in red and β-sheet in cyan and turns in green. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

graph shows a very good agreement for all residues, except for the 4–7 region, which exhibits a lower value for T_1/T_2, typical of protein regions presenting fast movements. The rest of the protein shows a moderate anisotropy, resulting in T_1/T_2 values that vary in the 8.5–6.3 interval. This behavior is the result of BA42 being a slightly prolate ellipsoid presenting an axially symmetric rotational diffusion tensor with an anisotropy ($D_{\parallel}/D_{\perp}$) of 1.23. The main axis of the diffusion tensor is almost parallel to α-helices I, IV, and V, and is almost perpendicular to α-helices II and III, as well as the HN vectors belonging to the central β-sheet. As a consequence, ^{15}N nuclei located in α-helices I, IV, and V show the highest T_1/T_2 values (around 8), due to a slower reorientation with respect to the magnetic field of the NH vectors belonging to these segments as compared to those pointing perpendicular to the main axis. This is the case for NHs belonging to the β-strands and α-helices II and III, which show a much lower T_1/T_2 value (around 6.3). The agreement between calculated and experimental T_1/T_2 ratios is a further validation of the structure, as they contain structural information similar to that of the residual dipolar couplings.[41] The first seven residues of BA42 constitute the only region presenting local flexibility, whereas the rest of the protein shows a marked rigidity. Notably, a long region of the protein, comprising residues 127–141, shows no defined secondary structure but at the same time the observed ^{15}N-relaxation rates [Fig. 3(A)] indicate the absence of flexibility.

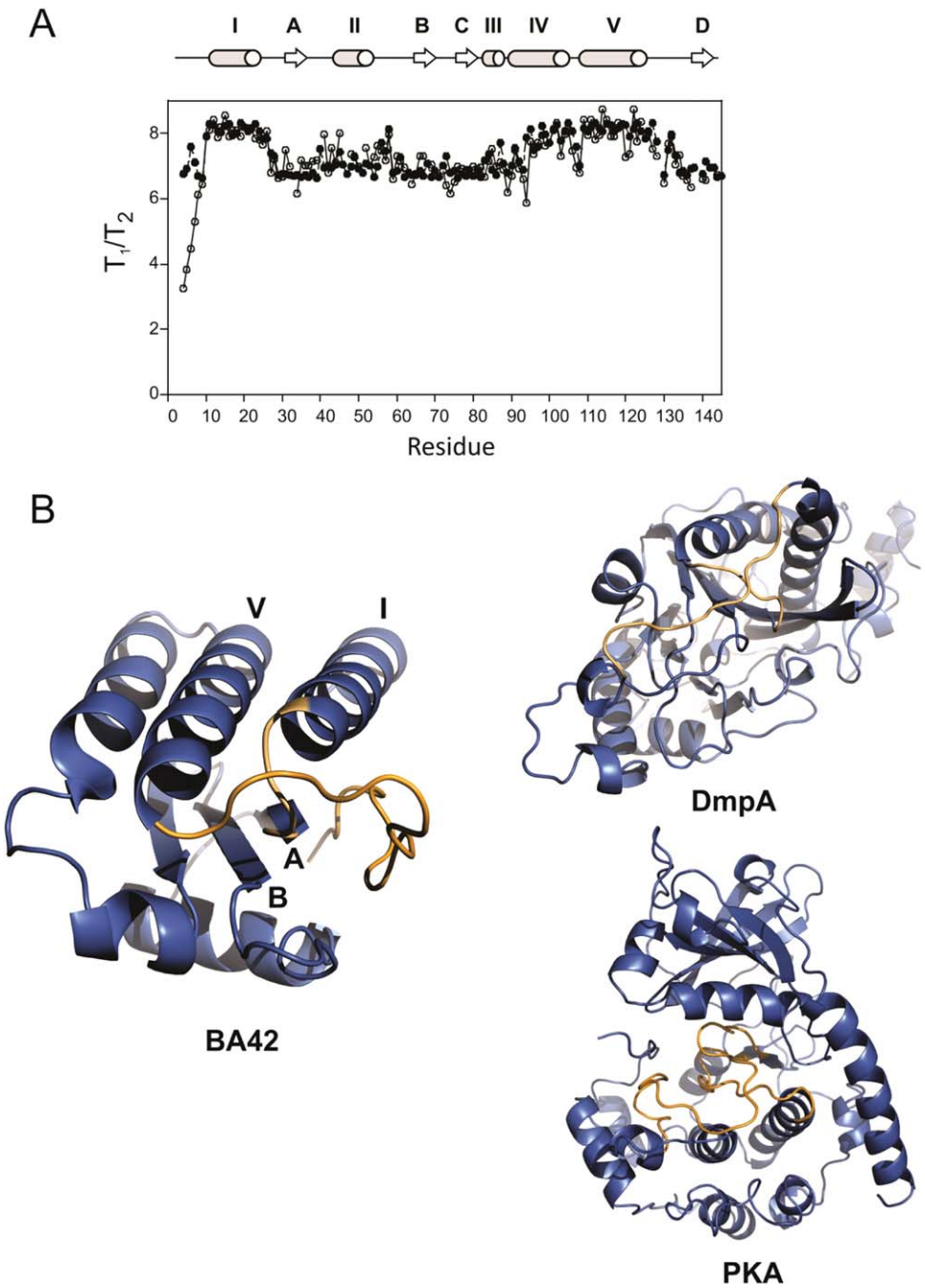

Figure 3

(**A**) Comparison of ^{15}N T_1/T_2 ratios of BA42, experimental values (open circles), and values calculated using the program HYDRONMR (black circles). Overhead, the secondary structure of BA42 is shown; the α-helices are represented by white arrows and the β-sheets by gray cylinders. (**B**) View of loop-crossing in the solution structure of BA42: the loop connecting the last α-helix (V) and the C-terminal β-strand (D) crosses over the loop between the N-terminal α-helix (I) and the first β-strand (A). On the right, two examples of structures having a pair of crossing loops, DmpA[42] (1B65) and cAMP-dependent protein kinase (PKA)[43] (1CDK), are shown. The loops are highlighted in orange. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The BA42 solution structure shows an infrequent structural feature: a pair of crossing loops

Careful examination of the solution structure of BA42 reveals the presence of a peculiar feature: the loop connecting the last α-helix (V) and the C-terminal β-strand (D) crosses over the loop between the N-terminal α-helix (I) and the first β-strand (A) [Fig. 3(B)]. This loop-crossing is very rare in protein structures and is thought to be energetically unfavorable.[44,45] No

particular interaction between the loops could be identified as a plausible explanation for the stabilization of this conformation. We can safely discard an alternative interpretation of some NOEs between residues of β-strands D and A, as being intermolecular instead of intramolecular. If this were the case, a swapped dimer, in which the C-terminal β-strand of one monomer pairs with the rest of the β-sheet of the other monomer and vice versa, would avoid the loop crossing in the BA42 structure. However, this alternative structure for BA42 is not compatible with the value of the hydrodynamic radius, and particularly, with the value of the correlation time calculated from ^{15}N relaxation rates and the T_2 value measured for ^{15}N nuclei. A dimeric structure of BA42, consisting of 290 residues and a MW of 33 kDa, would show a larger value for τ_c (in the order of 12 ns) and significantly shorter average ^{15}N T_2 values (around 50 ms) than those determined experimentally (8.0 ns and 96 ms, respectively).

The crystal structure of BA42 reveals a double metal binding site

In search of undiscovered features that may be responsible for the stabilization of the long C-terminal loop conformation packed against the structured core, we undertook the exploration of experimental conditions to crystallize the BA42 protein. High quality crystals of BA42 were obtained, which diffracted X-rays at a maximum resolution of 1.39 Å in the spacegroup $P2_12_12_1$. The crystal structure was solved by molecular replacement starting from the BA42 NMR model, as described before. The final coordinates show excellent stereochemistry, with more than 98% of the residues in the favored region of the Ramachandran plot (Table II). There is continuous density in the polypeptide chain without breaks. A few side chains corresponding to polar exposed residues could not be located in the electron density, and were therefore modeled only to their C^β.

The mother liquor used for crystallization did not contain divalent cations. However, clear electron density compatible with two divalent cations could be seen in the first cycles of refinement in the vicinity of residues E24, E30, D134, I135, D136, I138, and N140. Calcium was eventually assigned to this density after several concluding observations: (i) the Fourier maps yielded electron density levels compatible with calcium after refinement, (ii) the expected coordination geometry and binding distances for calcium were observed, (iii) there was lack of signal in the X-ray excitation scan ruling out the presence of transition metals, (iv) both calcium atoms and their surrounding water molecules show almost similar values for their B-factors after refinement.

The crystal and solution structures of BA42 are very similar with respect to the backbone conformation, with a C^α rmsd of 0.9 Å [Fig. 4(A)]. In accordance with the flexibility for the first seven residues observed in solution using ^{15}N relaxation measurements, no electron density was detected for the first four residues in the X-ray model. The crystal structure confirms that the loop V-D crosses over the loop I-A, but reveals a feature not detectable in the NMR solution experiments: the presence of two bound Ca^{2+} ions [represented by spheres in Fig. 4(A)]. The first Ca^{2+} ion is bound by the side chains of E30 and D134, and by the carbonyl oxygen of I135, with three water molecules completing the coordination sphere. The second Ca^{2+} ion interacts with the side chains of E24, E30, D136, and N140, with the carbonyl oxygen of I138 and with one water molecule [Fig. 4(B)]. This double metal binding site stabilizes the conformation of the crossing loops, as the two Ca^{2+} ions link the loop V-D (through residues D134, I135, D136, and N140) with two residues located at the edges of the loop crossing below: the first residue of the loop I-A (E24) and the first residue of β-strand A (E30). Additionally, two salt bridges between R21 (α-helix I) and D136 (loop V-D) and between E13 (α-helix I) and K144 (β-strand D) further stabilize the loop conformation.

Metal binding induces a significant conformational change in BA42

To prove that metal binding is crucial to stabilize the structure of BA42 with the C-terminal segment anchored to the αβα region, we treated a sample of ^{15}N-labeled protein with an excess of EDTA. The ^{1}H—^{15}N HSQC spectrum revealed a large number of peaks varying their position [Fig. 5(A)], indicative of a significant structural change. Elimination of EDTA and addition of an excess of Ca^{2+} restore an almost identical spectrum to that of the beginning (data not shown), indicating that metal binding is indeed the responsible for the structural change.

The HSQC spectrum of metal-free BA42 showed a reduced number of peaks, indicating that some amide groups are more solvent-exposed and/or they belong to regions with diminished conformational stability, leading to broader signals. Standard 3D experiments were conducted on a doubly labeled BA42 sample, leading to an almost complete assignment of the backbone nuclei of the metal-free state. Residues that were not assigned (because their cross peaks were not detected in the NMR experiments) were E40 (loop A-II), N53 (last residue of α-helix II), D57, K60, L61, N63 (loop II-B), R97 (α-helix IV), T141, and S143 (loop V-D). For the remaining residues, their chemical shift perturbation due to metal binding could be calculated, and is presented for the HN group in Figure 5(B). Regions exhibiting the largest chemical shift perturbation upon metal binding are the loop V-D and the β-strand D, followed by two regions that are in proximity to them in the structure of BA42: the loop I-A, the β-strand A, and the loop II-B. Most of

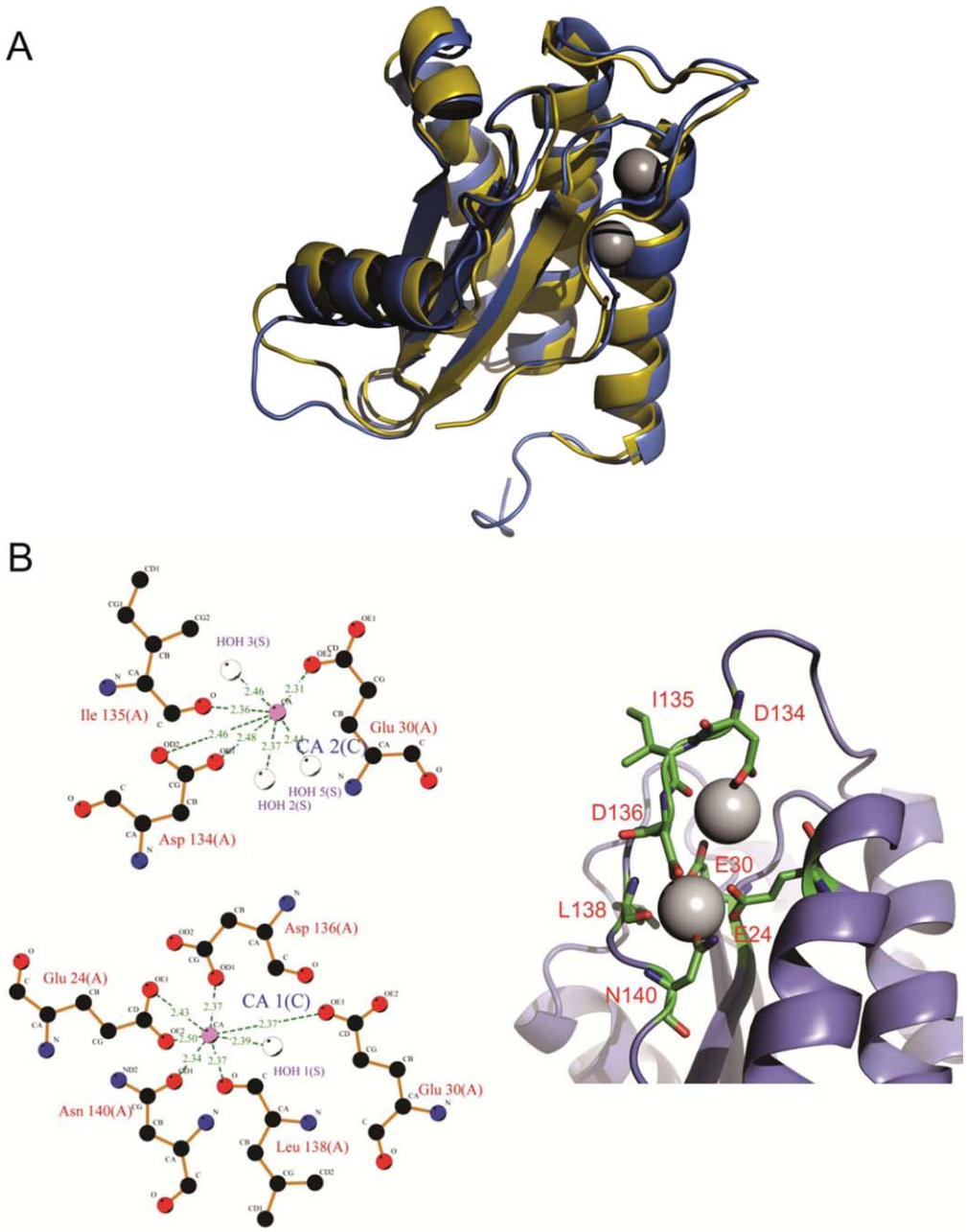

Figure 4

Crystal structure of BA42. (**A**) Schematic superposition of crystal structure (green) and solution structure (red). The bound Ca^{2+} atoms are displayed as green spheres. (**B**) Coordination of two bound Ca^{2+} ions. Left, the first Ca^{2+} ion is bound by the side chains of E30 and D134, by the carbonyl oxygen of I135 and interacts with three water molecules. The second Ca^{2+} ion is bound by the side chains of E24, E30, D136, and N140, with the carbonyl oxygen of I138 and interacts with one water molecule. Right, a detail showing the side chains interacting with the two bound Ca^{2+} atoms in the structure. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the residues whose cross-peaks were not detected belong to these regions, with the exception of E40 and R97. The amide proton of E40 is solvent-exposed in the metal-bound structure, showing already a low intensity signal in the $^{1}H—^{15}N$ HSQC spectrum of this form.

Despite the generalized change of the H—N chemical shifts, calculation of the chemical shift index using H^{α}, C^{α}, C^{β}, and C' chemical shift values of the BA42 metal-free form reveals that the secondary structure elements involve the same residues as in the metal-bound form

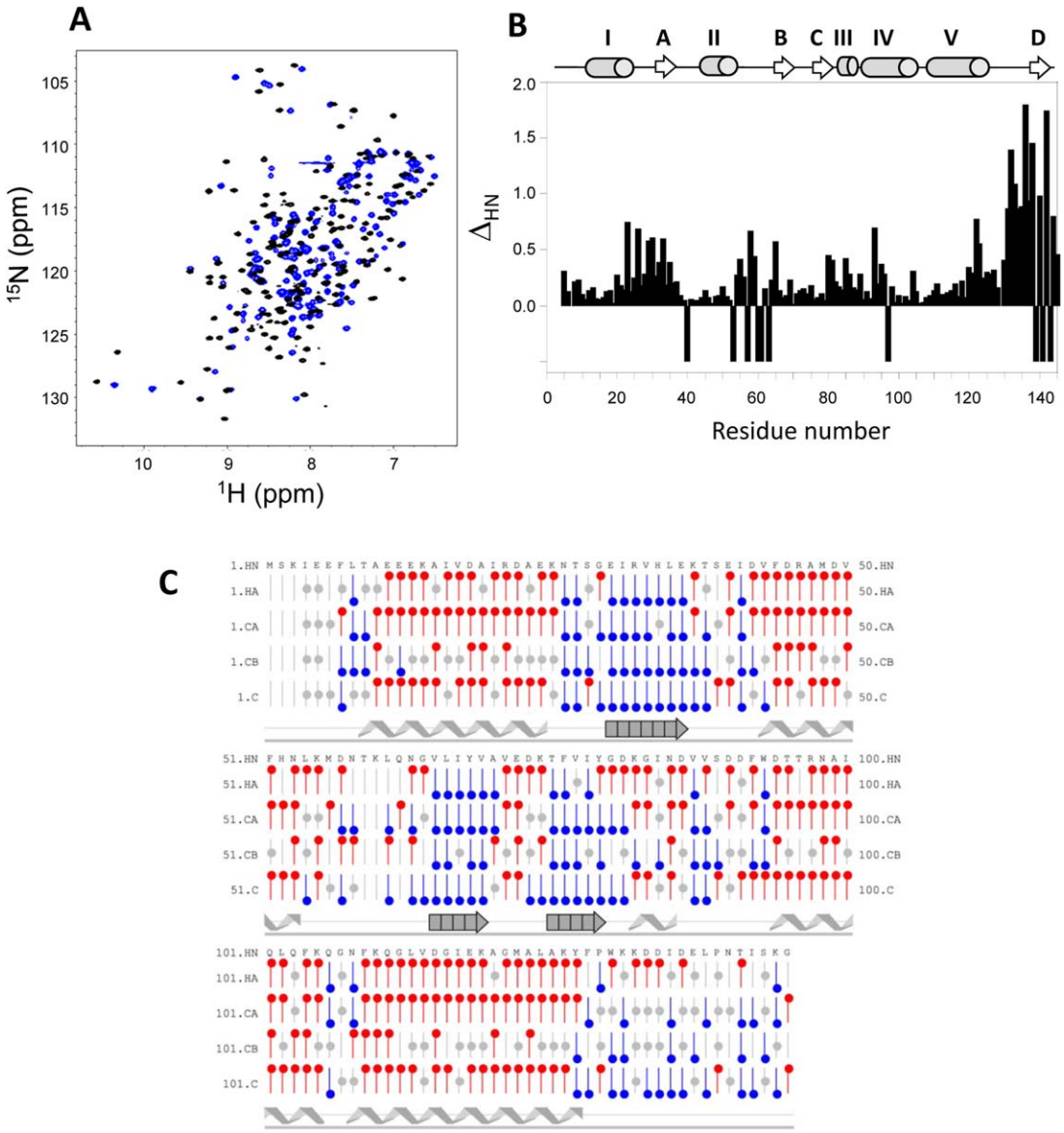

Figure 5

Conformational changes of BA42 inducted by calcium binding. (**A**) 1H—15N HSQC spectra of BA42. BA42 was incubated with an excess of EDTA (blue spectrum) and, after elimination of EDTA by dialysis, the protein was treated with 3 m*M* $CaCl_2$ (black spectrum). (**B**) HN chemical shift perturbation of BA42 in the presence or absence of Ca^{2+}. Not assigned amide nuclei are indicated with a negative index. The secondary structure of the metal-bound form is illustrated above. (**C**) Secondary structure prediction of metal-free BA42. The secondary structure was predicted from the chemical shift index calculated for Hα, Cα, Cβ, and C′. Red circles at the top, for shifts consistent with helical structures; blue circles at the bottom, for shifts consistent with sheet structures; and gray circles at the center, for intermediary shifts. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

[Fig. 5(C)]. All α-helices and β-strands remain unaltered after EDTA treatment, in particular β-strand A, which shows significant chemical shift perturbations.

To complete the picture about the conformational change provoked by the EDTA treatment, we measured the H—N NOE effect [Fig. 6(A)]. Values lower than 0.7 are indicative of flexibility in the ps time-scale. It is already evident that the region 131–145 is very flexible, exhibiting H—N NOE values in the range of 0.6 to −2.5. Figure 7(B) shows the difference between the measured H—N NOEs of the EDTA-treated form and the metal-bound form. Residues showing negatives ΔNOE are expected to be more rigid in the metal-bound form than in the metal-free form, and vice versa. Large and negative ΔNOE values were obtained for the C-terminal tail, indicating that the 131–145 segment dramatically increases its flexibility in the absence of Ca^{2+}. Interestingly, there is some evidence that the N-terminal region (4–7) is

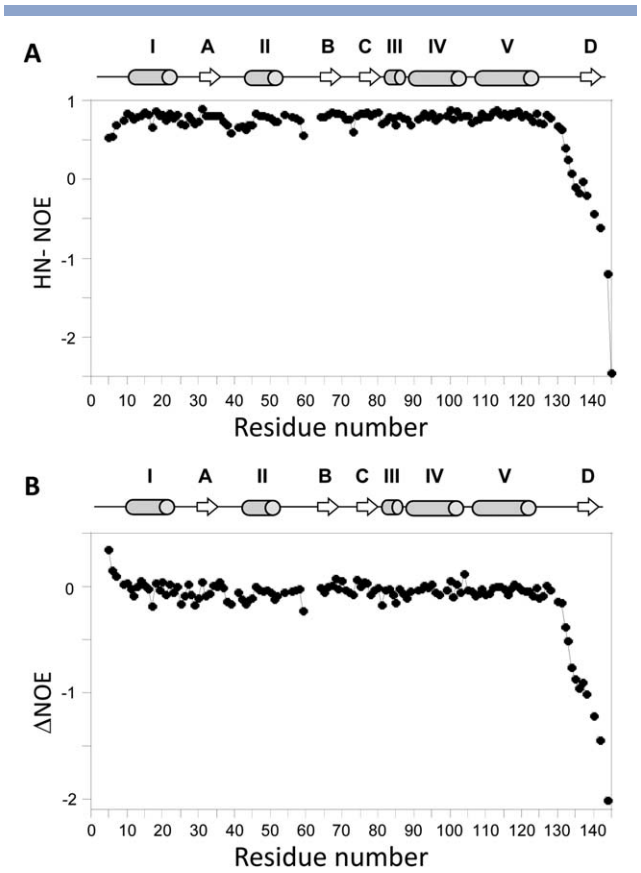

Figure 6

^{1}H—^{15}N heteronuclear NOE of BA42. (**A**) H—N NOEs measured for the EDTA-treated form of BA42. (**B**) Difference between the measured H—N NOEs for the EDTA-treated form and the metal-bound form. The secondary structure of the metal-bound form is illustrated above.

more structured after EDTA treatment. All these results together indicate that removal of the two metal ions from BA42 primarily changes the conformation of the C-tail, leading to a very flexible region comprising residues 131–145, without significantly weakening the stability and the position of the five α-helices and the three β-strands.

Calcium binding enhances the conformational stability of BA42

To assess whether the conformational change induced by Ca^{2+} binding affected the stability of the protein, we performed thermal denaturation experiments by CD. In a first experiment, we found that the far-UV spectra of BA42 in the metal-free form and in the presence of an excess of Ca^{2+} were almost identical; they are dominated by an α-helix content with two minima at 209 nm and 223 nm (Supporting Information material). This result is in line with the chemical shift analysis, indicating that the secondary structure is not altered by metal binding. The unfolding–folding process was reversible in both cases. To investigate the thermal unfolding, BA42 was incubated in the presence or absence of Ca^{2+} at different temperatures and the molar ellipticity at 223 nm was recorded. The thermal denaturation process presents a cooperative behavior where the melting temperature (T_m) increases with Ca^{2+} binding. The calculated T_m's for the metal-free BA42 and for the metal-bound BA42 were (47.6 ± 0.4) °C and (61.9 ± 0.4) °C, respectively.

BA42 shows a micromolar dissociation constant for Ca^{2+} binding

To evaluate the metal binding affinity, we titrated the metal-free form of BA42 with Ca^{2+}. We followed the binding by inspection of the high-field resonances in the range of 0.5 to −1.5 ppm in the 1D ^{1}H NMR spectrum. Figure 7(A) shows this spectral region during a titration experiment, evidencing the changes in chemical shifts of the methyl groups from V71, I78, V88, and I100, and one β-proton from M56. All these residues are highlighted in the metal-bound state structure in Figure 7(B). The fact that residues far away from the metal binding site, like V7, I78, I100, and V88, are sensitive to the variation in metal occupancy reflects a slight but significant change in the tertiary structure of the three-layered αβα structure. The large shielding effect observed for M56 H^{β} chemical shift in the metal-bound state is particularly indicative of the structural changes triggered by Ca^{2+} binding. It is provoked mainly by the proximity to L136, creating a highly hydrophobic environment for the side chain of M56. In the metal-free state, the chemical shift of the two β-protons of M56 move to 1.84 ppm, a value compatible with a solvent-exposed side chain.

During the course of the titration experiment, all peaks belonging to the metal-free protein decrease their intensities when the Ca^{2+} concentration rises, with a simultaneous appearance of the metal-bound form peaks. This indicates that the exchange kinetics is slow in the NMR-time scale. This behavior is observed even for peaks that show a small difference in chemical shift between the two forms, like I100, leading to an estimated upper limit for the exchange rate constant (k_{ex}) of 0.4 s^{-1} at 300 K. Such slow exchange kinetics reflects the large conformational change associated to metal binding.

We used the variation of the peak integrals as a function of the Ca^{2+} concentration to evaluate both the stoichiometry and the affinity of the interaction. Different samples, however, showed a lagged phase in which the addition of Ca^{2+} did not change the populations of the two forms [Fig. 7(C)]. Fitting the curve with a Hill equation varying the n coefficient yielded the best accordance for $n = 2$, indicating that there is a positive cooperative interaction between both cations. The total dissociation constant calculated from the fitting is 1826, leading an average K_d of 43 μM for each site, or equivalently, an affinity constant of $2.3 \times 10^4 M^{-1}$. This relatively high to

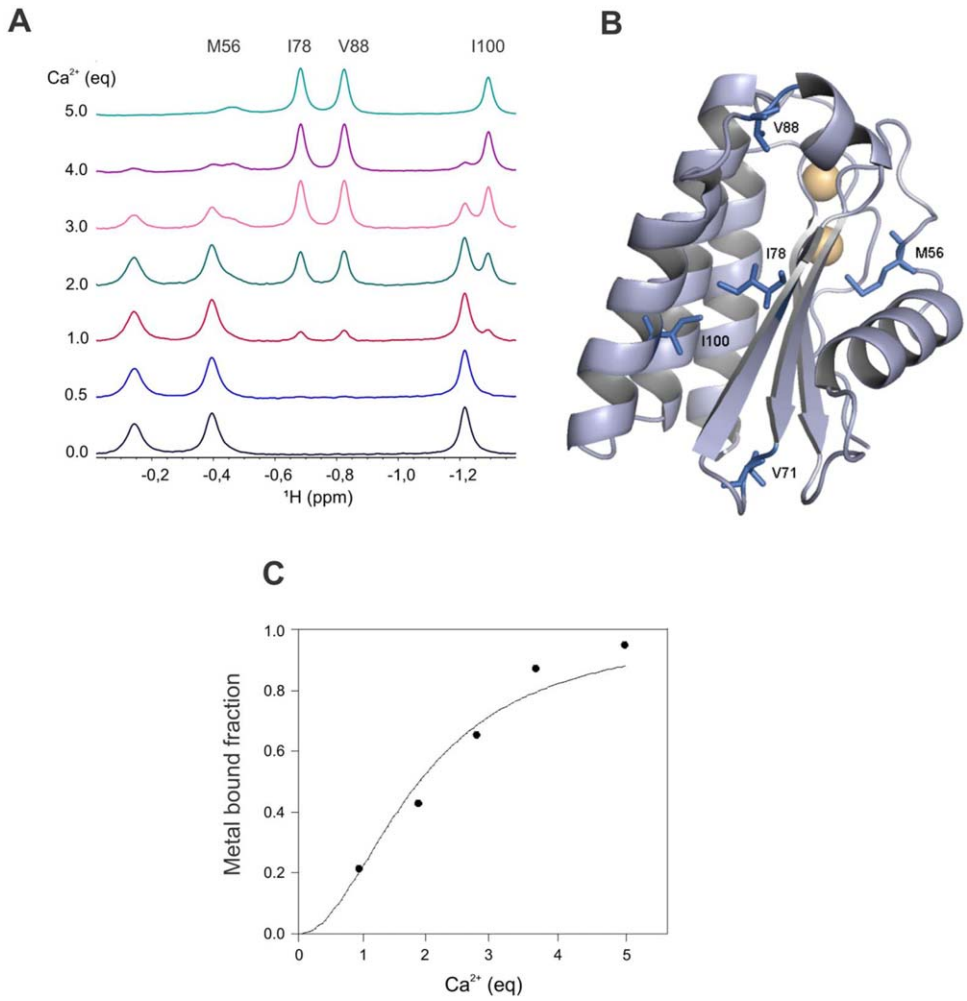

Figure 7

Metal binding affinity. (**A**) Titration of metal-free BA42 with Ca^{2+} followed by 1D 1H NMR spectra in the range of 0.5 to −1.5 ppm. The methyl groups from V71, I78, V88, and I100, and one β-proton from M56 are indicated. (**B**) Ribbon view of metal-bound structure evidencing V71, I78, V88, I100, and M56 residues. (**C**) Variation of the peak integral of 1D 1H NMR spectra as a function of the Ca^{2+} concentration. The curve was fitted using the Hill equation (full line) giving an *n* coefficient of 2. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

moderate affinity is compatible with the fact that in the samples used for NMR and X-ray crystallography, no trace of the metal-free protein was observed. It can be thus concluded that the protein was able to pick up two Ca^{2+} ions from the culture medium, without the need of metal addition.

PF04536 members are the only structural homologs of BA42

A search for structural homologs of BA42 with the program Dali[46] showed that the closest structures are those of the other three PF04536 members [Fig. 8(A)], confirming that this type of α/β architecture is a distinctive feature of this family. The best matches are the entries PG0361 (*Z*-score 12.0, C^{α} rmsd = 2.9 Å for 124 aligned residues with 15% sequence identity), CG2496 (11.6, 2.7 Å, 114, 15%), and AtTLP18.3 (11.1, 2.8 Å, 119, 12%). AtTLP18.3 shows two additional α-helices (I and IV) and shares with BA42 the presence of a structural element corresponding to α-helix III (α-helix V). One distinct feature of BA42 is the topology of its β-sheet: AtTLP18.3, PG0361, and CG2496 exhibit an A(↑)B(↑)C(↑)D(↓) topology, in contrast to the D(↑)A(↑)B(↑)C(↓) topology of BA42 [Fig. 8(B)]. In general, the long C-terminal tail of BA42 has a similar position to that of the N-terminus in the other three structures, resulting in a circular permutation of the fold. A second characterizing feature of BA42 is the presence of the two metal binding sites, which favors the positioning of the C-tail to complete the fourth strand of the β-sheet.

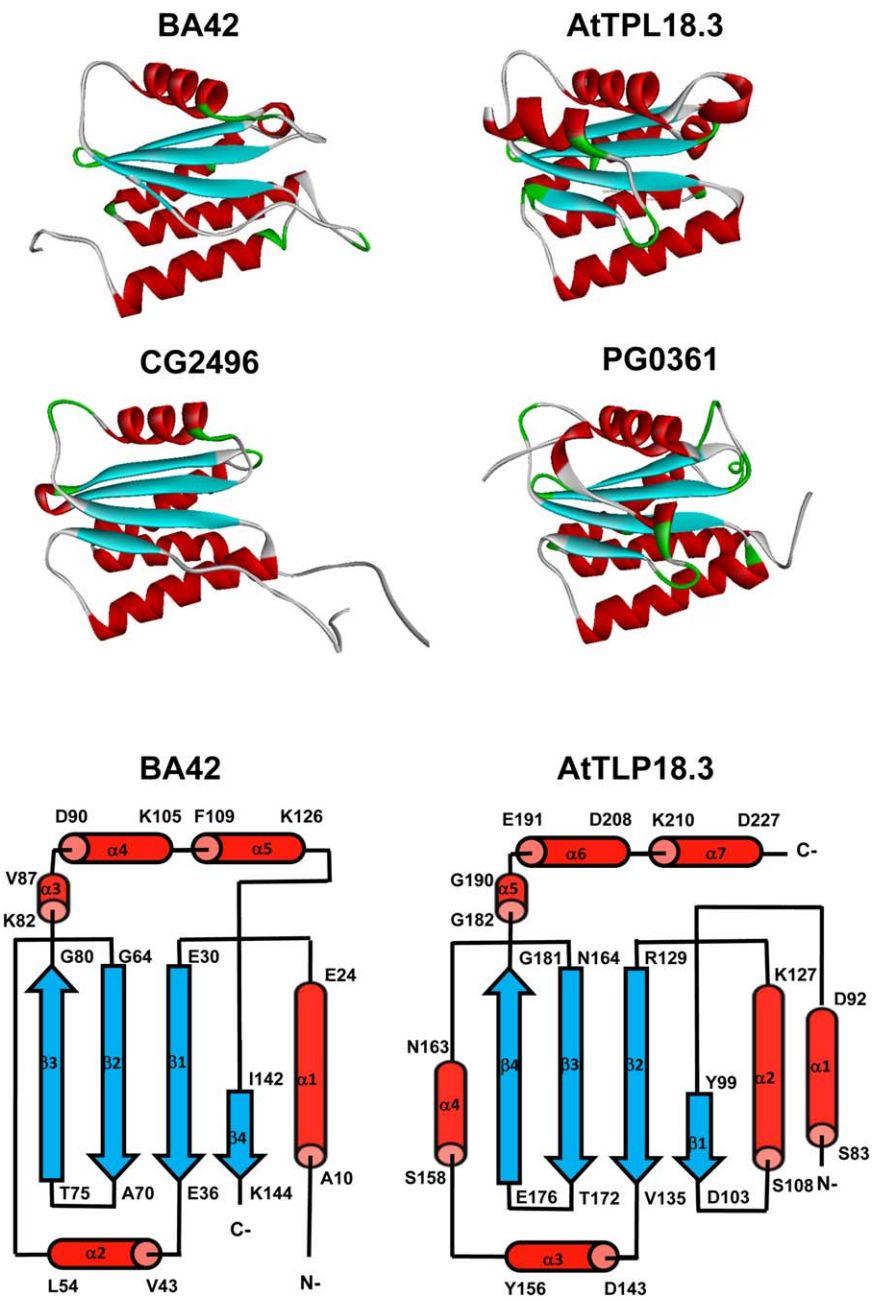

Figure 8

(**A**) Known structures of the PF04536 family members. They include AtTPL18.3 from *Arabidopsis thaliana* (3PTJ), CG2496 from *Corynebacterium glutamicum*, (2KPT), and PG0361 from *Porphyromonas gingivalis* (2KW7). (**B**) Topology of BA42 and AtTLP18.3 showing the different disposition of the four strands forming the central β-sheet.

DISCUSSION

The structural and functional knowledge of the PF04536 family are slowly growing. The peculiar αβα sandwich structure shown by these protein domains has no close structural homologs outside this protein family. The BA42 protein, however, exhibits two distinctive features. First, BA42 is the only protein, among those for which a structure was determined, constituted solely by a TPM domain; the other three proteins are multi-domain. A Pfam analysis of the sequence of the two bacterial proteins CG2496 (684 residues) and PG0361 (435 residues) shows the probable presence of a signal peptide in the first 30–40 residues, and other putative trans-membrane

domains toward the C-termini of these proteins. The *Arabidopsis thaliana* acid phosphatase AtTLP18.3 (285 residues) shows both a chloroplast import peptide and a thylakoid lumen import peptide at its N-terminus, and a trans-membrane region at its C-terminus.[11] Very recently, a TPM domain was also identified as part of MOLO-1, a positive regulator of levamisole-sensitive acetylcholine receptor at the *Caenorhabditis elegans* neuromuscular junction.[47] As in the other cases, the TPM domain of MOLO-1 constitutes a single extracellular globular domain of a protein containing a one-pass transmembrane segment. In all these proteins, the TPM domain is linked to proteins that are imported and anchored to membranes. The existence of a single TPM domain constituting the entire protein is a characteristic that seems to be present only in members of the family Flavobacteriaceae. Second, the topology of the BA42 fold represents a yet unknown topological variant of the TPM domain family and is related to the topology of the common core of the other three structures by a circular permutation. The BA42 fold can be converted to the "common" fold observed for CG2496, PG0361, and AtTLP18.3 by connecting the N- and C-termini by a linker and creating the new termini by cleaving the loop between the last α-helix (V) and the C-terminal β-strand (D).

The completion of the four-stranded β-sheet either by an N-terminal β-strand (in the three already known TPM domain structures) or by a C-terminal β-strand (BA42) gives rise to rarely observed topological features. In the structures of CG2496, PG0361, and AtTLP18.3 the βαβ unit, formed by the first N-terminal β-strand, the connecting α-helix and the second β-strand, is left-handed. This kind of βαβ supersecondary structure unit containing a two-stranded parallel β-sheet and a crossover α-helix, is right-handed in ~98.5% of the structures in the non-redundant protein database.[48] The circular permutation in the BA42 fold breaks the left-handed unit, but creates an even rarer topological feature, a pair of crossing loops as described before. In the context of the BA42 structure, this conformation is stabilized by the presence of a double metal binding site. The residues constituting these metal binding sites are highly conserved in all BA42 homologs (see Fig. 1) suggesting that this unusual and distinctive characteristic is shared by all these proteins.

Crossing connections are unfavorable because the crossing both buries and dehydrates a loop peptide group (which costs around 5 kcal mol^{-1}),[44] or demands additional loop bending to avoid this dehydration, which also costs a few kcal mol^{-1}.[45] In more recent works, however, it is recognized that the rule "no loop crossovers" finds an increasing number of violations in the folds deposited in the PDB, although it is still present in a reduced number of proteins.[49] Examples of protein structures containing crossing-loops are the DOM-fold (a consensus structure of the domains in DmpA, L-aminopeptidase D-Ala-esterase/amidase, OAT, ornithine acetyltransferase, and MocoBD, molybdenum cofactor-binding domain)[42] [Fig. 3(B)] and the *sodC*-encoded *Mycobacterium tuberculosis* superoxide dismutase (SOD).[50] The most relevant example of the presence of this crossover connection is the kinase superfamily[43,51] where it is considered a hallmark. Both the structure of type IIβ phosphatidylinositol phosphate kinase (PIPK) and serine/threonine and tyrosine kinases (PKs) share a pair of crossing loops in their C-terminal domains [Fig. 3(B)].[43] In particular, it was observed that the "crossing loops" structure is formed in the catalytic region of PKs, where a magnesium ion is fundamental for ATP interaction and the phosphotransfer reaction.[51,52] To our knowledge, however, there is no example of a crossover connection stabilized by direct formation of a metal binding site, as we found in BA42.

In response to low temperatures, psychrophiles use diverse strategies for adaptation.[53–55] In most organisms, the conformational flexibility of proteins is adjusted, resulting in a loose structure characterized by a low conformational stability.[56,57] This feature is preponderantly observed in the metal-free form of BA42. In fact, an increase of 14°C in the melting temperature is caused by metal binding, revealing a strong difference in the conformational stability between the two forms. The transition between the native and unfolded states occurs by a largely cooperative event, which is in line with the proposed model of folding funnels for psychrophilic enzymes.[58] As in the case of other cold-adapted proteins studied so far, the unfolding mechanism is reversible, both in the absence or presence of metals.[48]

Probably related to this change in thermal stability, the overall structural dynamics of the two forms of BA42 are also different. The metal-bound form is characterized by a notable rigidity, with the exception of the first seven residues. Without the metal, BA42 undergoes a transition to a more flexible structure, presenting a long C-terminal tail of 17 residues largely disordered. Most of the HN-NOEs measured for the metal-free form in the structured core showed a slightly smaller value with respect to the metal-bound form, indicating that overall the whole core is less compact. In this respect, BA42 shows a flexibility that is modulated by the metal concentration. This dual behavior may reflect a differential need in the two states of the protein, as postulated for other psychrophilic proteins.[58] It can be hypothesized that the flexible form of metal-free BA42, in particular around the metal binding domain, may favor the interaction with the metals at low temperature by reducing the energy cost of induced-fit mechanisms. Conversely, it is possible that the rigidity of the metal bound form can be used to improve binding to a yet unknown molecular target, by increasing the affinity constant and the selectivity as a consequence of a

reduced mobility, as already observed for the heat-stable domain of the psychrophilic phosphoglycerate kinase.[59]

Finally, the averaged value of the metal binding affinity constant is $2.3 \times 10^4 M^{-1}$ for each site. This value is remarkably similar to that of the lowest affinity calcium-binding site of calmodulin measured at the same buffer, pH, and similar ionic strength and temperature ($2.5 \times 10^4 M^{-1}$).[60] This suggests a possible role of BA42 in sensing the Ca^{2+} concentration, which triggers a large structural rearrangement that stabilizes the conformation and possible changes its activity. Further experiments are in progress to determine the presence of an enzymatic activity in either of the two forms of BA42, which can help in revealing its possible role in the context of the life-cycle of *B. argentinensis*.

ACKNOWLEDGMENTS

The authors acknowledge access to the PROXIMA 1 beamline at the SOLEIL Synchrotron (France). The authors are grateful to the Argentinian Ministry of Science (MINCyT) for travel support and to Dr. Federico Iacovelli and Prof. Mattia Falconi, Department of Biology, University of Rome "Tor Vergata," for conducting the final minimizations of the NMR structures in water.

REFERENCES

1. Yi H, Oh HM, Lee JH, Kim SJ, Chun J. *Flavobacterium antarcticum* sp. nov., a novel psychrotolerant bacterium isolated from the Antarctic. Int J Syst Evol Microbiol 2005;55:637–641.
2. Bercovich A, Vazquez SC, Yankilevich P, Coria SH, Foti M, Hernández E, Vidal A, Ruberto L, Melo C, Marenssi S, Criscuolo M, Memoli M, Arguelles M, Mac Cormack WP. *Bizionia argentinensis* sp. nov., isolated from marine water in Antarctica. Int J Syst Evol Microbiol 2008;58:2363–2367.
3. Lanzarotti E, Pellizza L, Bercovich, Foti M, Coria S, Vazquez S, Ruberto L, Hernández E, Dias R, Mac Cormack W, Cicero DO, Smal C, Nicolas M, Ribeiro Vasconcelos A, Marti M, Turjanski A. Draft genome sequence of Bizionia argentinensis, isolated from Antarctic surface water. J Bacteriol 2011;193:6797–6798.
4. Levitt M. Nature of the protein universe. Proc Natl Acad Sci USA 2009;106:11079–11084.
5. Kamil Khafizov, Carlos Madrid-Aliste, Steven C. Almo, Andras Fisera. Trends in structural coverage of the protein universe and the impact of the protein structure initiative. Proc Natl Acad Sci USA 2014;111:3733–3738.
6. Terwilliger TC. The success of structural genomics. J Struct Funct Genomics 2011;12:43–44.
7. Fajardo JE, Fiser A. Protein structure-based prediction of catalytic residues. BMC Bioinformatics 2013;14:63.
8. Stein A, Céol A, Aloy P. 3did: identification and classification of domain-based interactions of known three-dimensional structure. Nucleic Acids Res 2011;39(Database issue):D718–D723.
9. Smal C, Aran M, Lanzarotti E, Papouchado M, Foti M, M. Marti A, Coria SH, Vazquez SC, Bercovich A, Mac Cormack WP, Turjanski AG, Gallo M, Cicero DO. ^{1}H, ^{15}N and ^{13}C chemical shift assignments of the BA42 protein of the psychrophilic bacteria *Bizionia argentinensis* sp. nov. Biomol NMR Assign 2012;6:181–183.
10. Eletsky A, Acton TB, Xiao R, Everett JK, Montelione GT, Szyperski T. Solution NMR structures reveal a distinct architecture and provide first structures for protein domain family PF04536. J Struct Funct Genomics 2012;13:9–14.
11. Wu H-Y, Liu M-S, Lin T-P, Cheng Y-S. Structural and functional assays of AtTLP18.3 identify its novel acid phosphatase activity in thylakoid lumen. Plan Physiol 2011;157:1015–1025.
12. Sirpiö S, Allahverdiyeva Y, Suorsa M, Paakkarinen V, Vainonen J, Battchikova N, Aro EM. TLP18.3, a novel thylakoid lumen protein regulating photosystem II repair cycle. Biochem J 2007;406:415–425.
13. Cicero DO, Contessa GM, Paci M, Bazzo R. HACACO revisited: residual dipolar coupling measurements and resonance assignments in proteins. J Magn Reson 2006;180:222–228.
14. Ottiger M, Delaglio F, Bax A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson 1998;131:373–378.
15. Eliseo T, Ragona L, Catalano M, Assfalg M, Paci M, Zetta L, Molinari H, Cicero DO. Structural and dynamic determinants of ligand binding in the ternary complex of chicken liver bile acid binding protein with two identical bile salts revealed by NMR. Biochemistry 2007;46:12557–12567.
16. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
17. Johnson BA. Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 2004;278:313–352.
18. Wu DH, Chen AD, Johnson CS. An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. J Magn Reson A 1995;115:260–264.
19. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ. Hydrodynamic radii of native and denaturated proteins measured by pulse field gradient NMR techniques. Biochemistry 1999;38:16424–16431.
20. Kay LE, Torchia DA, Bax A. Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry 1989;28:8972–8979.
21. Barbato G, Ikura M, Kay LE, Pastor RW, Bax A. Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. Biochemistry 1992;31:5269–5278.
22. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. J Magn Reson 2003;160:65–73.
23. O'Donoghue SI, King GF, Nigles M. Calculation of symmetric multimer structures from NMR data using a priori knowledge of the monomer structure, co-monomer restraints, and interface mapping: the case of leucine zippers. J Biomol NMR 1996;8:193–206.
24. Clore GM, Gronenborn AM. New methods of structure refinement for macromolecular structure determination by NMR. Proc Natl Acad Sci USA 1998;95:5891–5898.
25. Tjandra N, Bax A. Direct measurement of distances and angles in biomolecules by NMR in dilute liquid crystalline medium. Science 1997;278:1111–1114.
26. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 2009;44:213–223.
27. Vriend G. WHAT IF: a molecular modeling and drug design program. J Mol Graph 1990;8:52–56.
28. Losonczi JA, Andrec M, Fischer MW, Prestegard JH. Order matrix analysis of residual dipolar couplings using singular value decomposition. J Magn Reson 1999;138:334–342.
29. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 2003;24:1999–2012.
30. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 2008;4:435–447.

31. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
32. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr 2010;66(Pt 2):125–132.
33. Navaza JA. AMoRe: an automated package for molecular replacement. Acta Crystallogr Sect A 1994;50:157–163.
34. Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vonrhein C, Womack TO. BUSTER version X.Y.Z. Cambridge, United Kingdom: Global Phasing, Ltd.; 2011.
35. Chen VB, Arendall WB, Headd JJ, III, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 2010;66:12–21.
36. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010;66:486–501.
37. Lin M, Larive CK. Detection of insulin aggregates with pulsed-field gradient nuclear magnetic resonance spectroscopy. Anal Biochem 1995;229:214–220.
38. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. Biochemistry 1999;38:16424–16431.
39. Daragan VA, Mayo KH. Motional model analyses of protein and peptide dynamics using ^{13}C and ^{15}N NMR relaxation. Prog Nucl Magn Reson Spectrosc 1997;31:63–105.
40. García de la Torre J, Huertas ML, Carrasco B. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. J Magn Reson 2000;147: 138–146.
41. Ryabov Y, Clore GM. Direct use of ^{15}N relaxation rates as experimental restraints on molecular shape and orientation for docking of protein-protein complexes. J Am Chem Soc 2010;132:5987–5989.
42. Cheng H, Grishin NV. DOM-fold: a structure with crossing loops found in DmpA, ornithine acetyltransferase, and molybdenum cofactor-binding domain. Protein Sci 2005;14:1902–1920.
43. Grishin NV. Phosphatidylinositol phosphate kinase: a link between protein kinase and glutathione synthase folds. J Mol Biol 1999;291: 239–247.
44. Finkelstein AV, Ptitsyn OB. Why do globular proteins fit the limited set of folding patterns? Prog Biophys Mol Biol 1987;50:171–190.
45. Finkelstein AV, Gutun AM, Badretdinov A. Why are the same protein folds used to perform different functions? FEBS Lett 1993;325: 23–28.
46. Holm L, Rosenström P. Dali server: conservation mapping in 3D. Nucl Acids Res 2012:38(suppl 2):W545–W549.
47. Boulin T, Rapti G, Briseño-Roa L, Stigloher C, Richmond JE, Paoletti P, Bessereau J-L. Positive modulation of a Cys-loop acetylcholine receptor by an auxiliary transmembrane subunit. Nat Neurosci 2012;15:1374–1381.
48. Cole BJ, Bystroff C. Alpha helical crossovers favor right-handed supersecondary structures by kinetic trapping: the phone cord effect in protein folding. Protein Sci 2009;18:1602–1608.
49. Grainger B, Sadowski ML, Tayler WR. Re-evaluating the "rules" of protein topology. J Comp Biol 2010;17:1371–1384.
50. Spagnolo L, Törő I, D'Orazio M, O'Neill P, Pederse JZ, Carugo O, Rotilio G, Battistoni A, Djinović-Carugo K. Unique features of the sodC-encoded superoxide dismutase from Mycobacterium tuberculosis, a fully functional copper-containing enzyme lacking zinc in the active site. J Biol Chem 2004;279:33447–33455.
51. Scheef ED, Bourne PE. Structural evolution of the protein kinase-like superfamily. PLOS Comp Biol 2005;1:359–381.
52. Taylor SS, Radzio-Andzelm E. Three protein kinase structures define a common motif. Structure 1994;2:345–355.
53. Russell NJ. Psychrophilic bacteria—molecular adaptations of membrane lipids. Comp Biochem Physiol A Physiol 1997;118:489–493.
54. Galtier N, Lobry JR. Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. J Mol Evol 1997;44:632–636.
55. Khachane AN, Timmis KN, dos Santos VA. Uracil content of 16S rRNA of thermophilic and psychrophilic prokaryotes correlates inversely with their optimal growth temperatures. Nucleic Acids Res 2005;33:4016–4022.
56. D'Amico S, Sohier JS, Feller G. Kinetics and energetics of ligand binding determined by microcalorimetry: insights into active site mobility in a psychrophilic alpha-amylase. J Mol Biol 2006;358: 1296–1304.
57. Feller G. Life at low temperatures: is disorder the driving force? Extremophiles 2007;11:211–216.
58. D'Amico S, Claverie P, Collins T, Georlette D, Gratia E, Hoyoux A, Meuwis M-A, Feeler G, Gerady C. Molecular basis of cold adaptation. Philos Trans R Soc Lond B 2002;357:917–925.
59. Bentahir M, Feller G, Aittaleb M, Lamotte-Brasseur J, Himri T, Chessa JP, Gerday C. Structural, kinetic, and calorimetric characterization of the cold-active phosphoglycerate kinase from the antarctic Pseudomonas sp. TACII18. J Biol Chem 2000;275:11147–11153.
60. Linse S, Helmersson A, Forsen S. Calcium binding to calmodulin and its globular domains. J Biol Chem 1991;266:8050–8054.