The CBS Domain Protein MJ0729 of *Methanocaldococcus jannaschii* Is a Thermostable Protein with a pH-Dependent Self-Oligomerization†

Luis Alfonso Martínez-Cruz,*,‡ José A. Encinar,§ Danel Kortazar,‡ Jesús Prieto,|| Javier Gómez,§ Pablo Fernández-Millán,‡ María Lucas,‡ Egoitz Astigarraga Arribas,⊥ José Andrés Fernández,⊥ María Luz Martínez-Chantar,# José M. Mato,# and José Luis Neira*,§,+

Unidad de Biología Estructural, CIC bioGUNE, Parque Tecnológico de Vizcaya, Ed. 800, 48160 Derio, Bizkaia, Spain, Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202 Elche (Alicante), Spain, Structural Biology and Biocomputing Programme, Centro Nacional de Investigaciones Oncológicas (CNIO), 28007 Madrid, Spain, Departamento de Química-Física, Universidad del País Vasco UPV-EHU, Lejona, Bizkaia, Spain, Unidad de Metabolómica, CIC bioGUNE, Parque Tecnológico de Vizcaya, Ed. 801, 48160 Derio, Bizkaia, Spain, and Instituto de Biocomputación y Física de los sistemas complejos, Zaragoza, Spain

Received October 13, 2008; Revised Manuscript Received December 19, 2008

ABSTRACT: CBS domains are small protein motifs, usually associated in tandems, that are involved in binding to adenosyl groups. In humans, several genetic diseases have been associated with mutations in CBS domains, and then, they can be considered as promising targets for the rational design of new drugs. However, there are no structural studies describing their oligomerization states, conformational preferences, and stability. In this work, the oligomerization state, the stability, and conformational properties of the CBS domain protein MJ0729 from *Methanocaldococcus jannaschii* were explored by using a combination of hydrodynamic (namely, ultracentrifugation, DLS, DOSY-NMR, and gel filtration) and spectroscopic techniques (fluorescence, circular dichroism, NMR, and FTIR). The results indicate that the protein had a pH-dependent oligomerization equilibrium: at pH 7, the dominant species is a dimer, where each monomer is a two-CBS domain protein, and at pH 4.5−4.8, the dominant species is a tetramer, with an oblong shape, as shown by X-ray. Deconvolution of the FTIR spectra indicates that the monomer at physiological pH has 26% α-helical structure and 17% β-sheet, with most of the structure disordered. These results are similar to the percentages of secondary structure of the monomer in the resolved tetrameric X-ray structure (21% of α-helical structure and 7% of β-sheet). At pH 2.5, there was a decrease in the level of secondary structure of the monomer, and formation of intermolecular hydrogen bonds, as shown by FTIR, suggesting the presence of high-molecular weight species. The physiological dimeric species is thermal and chemically very stable with a thermal midpoint of ∼99 °C, as shown by both DSC and FTIR; the GdmCl chemical midpoint of the dimeric species occurs in a single step and was greater than 4 M.

The cystathionine β-synthase (CBS)[1] domain proteins comprise a large superfamily of evolutionarily conserved proteins that are present in all kingdoms of life (*1, 2*). CBS domains are 60-residue long motifs (IPR000644, InterPro database) that usually occur as tandem pairs (two or four copies) either in stand-alone proteins or fused to different protein domains. They were first described in inosine monophosphate dehydrogenase (*2*). Although their function is unknown, their importance is underlined by the fact that mutations within CBS domains cause several hereditary diseases in humans, such as homocystinuria, autosomic retinitis pigmentosa, myotonia congenital, idyopatic epilepsy, or hypercalciuric nephrolytiasis, among others. Thus, they can be considered as promising targets for the development of novel drugs (*3*).

Comparison among CBS domains of isofunctional proteins from different species shows a highly conserved fold despite the low degree of sequence similarity (*2, 4*). In all these

† This work was supported by Projects SAF2008-05742-C02-01 and CSD-2008-00005 from the Spanish Ministerio de Ciencia e Innovación, and FIPSE (36557/06) to J.L.N., and by program grants from the Basque Government (ETORTEK IE05-147 and E07-202), Diputación Foral de Bizkaia (Exp. 7/13/08/2006/11 and 7/13/08/2005/14), and the Spanish Ministry of Education (SAF2005-00855) to L.A.M.-C. Support from CIC bioGUNE for predoctoral and postdoctoral fellowships is gratefully acknowledged.

* To whom correspondence should be addressed. L.A.M.-C.: Unidad de Biología Estructural, CIC bioGUNE, Parque Tecnológico de Bizkaia, Ed. 800, 48160 Derio, Vizcaya, Spain; telephone, +34 944061320; fax, +34 944061301; e-mail, amartinez@cicbiogune.es. J.L.N.: Instituto de Biología Molecular y Celular, Edificio Torregaitán, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202 Elche (Alicante), Spain; telephone, +34 966658459; fax, +34 966658758; e-mail, jlneira@umh.es.

‡ Unidad de Biología Estructural, CIC bioGUNE.

§ Universidad Miguel Hernández.

|| Centro Nacional de Investigaciones Oncológicas (CNIO).

⊥ Universidad del País Vasco UPV-EHU.

Unidad de Metabolómica, CIC bioGUNE.

+ Instituto de Biocomputación y Física de los sistemas complejos.

[1] Abbreviations: AU, analytical ultracentrifugation; β-ME, β-mercaptoethanol; CBS, cystathionine β-synthase; CD, circular dichroism; DLS, dynamic light scattering; DOSY, diffusion-ordered spectroscopy; FTIR, Fourier transform infrared spectroscopy; GdmCl, guanidinium hydrochloride; IR, infrared spectroscopy; NMR, nuclear magnetic resonance spectroscopy; T_m, thermal denaturation midpoint; UV, ultraviolet.

10.1021/bi801920r CCC: $40.75
Published on Web 03/06/2009

structures, two CBS domains associate to form a compact fold known as a Bateman domain (*5*). The cleft between both domains is the binding site for adenosyl groups; both head-to-head and head-to-tail associations of the Bateman domains have been observed for proteins from prokaryotic and eukariotic organisms, though the forces directing their final assembly are not known (*6*). Recently, the crystal structures of the complexes of several CBS domain-containing proteins with nucleotides, such as CLC5-ADP/ATP (*7*) and AMPK-ATP/AMP/ADP/ZMP (*8*−*12*), have been reported. These studies have provided crucial information that aids in understanding the architecture in binding to nucleotides.

CBS domains are unusually abundant in archaea, although scarce information about their function has been reported. Therefore, organisms such as the hyperthermophile *Methanocaldococcus jannaschii* (*13*) offer excellent models for the characterization of the adenosyl binding site of CBS domains. The genome of *M. jannaschii* encodes 15 CBS domain proteins (www.tigr.org), which differ significantly in their composition and presumably in their abilities to bind to different ligands (Figure 1A). A close examination of their amino acid sequences reveals the presence of two different groups: (i) one with very short sequences (from 124 to 186 residues) and not fused to other domains (as in the MJ0729 protein) (Figure 1B) and (ii) another with longer amino acid sequences which are fused to other protein motifs (from 214 to 509 residues) (Figure 1B). Only two of these latter proteins (so-called MJ1004 and MJ1225) contain four CBS domains in tandem, as in the γ-subunit of 5′-AMP-activated protein kinase (AMPK) and its homologues (*8*). The open reading frame of gene *mj0729* (UniProtKB/Swiss-Prot entry Q58139) encodes a polypeptide chain of 124 amino acids with a molecular mass of 14.303 kDa. Its sequence is formed by a CBS domain pair (CBS1, residues 13−60; CBS2, residues 73−122) (http://smart.embl-heidelberg.de/). Although MJ0729 is currently annotated as an uncharacterized hypothetical protein, it appears close to the genes encoding an iron−sulfur flavoprotein (MJ0731), very similar to a homologue found in *Archaeoglobus fulgidus* (*14*). In fact, on the basis of gene arrangements with the genome of *A. fulgidus*, it has been suggested that MJ0729 might be involved in the regulation of electron transport systems.

On the other hand, since MJ0729 is a homoligomer, as shown by the X-ray structure at pH 4.5−4.8 (*15*) (Figure 1C), analysis of its oligomerization and stability can provide insights into the molecular basis of protein assembly, regulation, and evolution. In this work, we describe the pH-dependent oligomerization and conformational properties of MJ0729 by using an array of hydrodynamic and spectroscopic techniques. We observe that MJ0729 had a pH-dependent oligomerization equilibrium, involving tetramers, trimers, dimers, and probably monomers at basic pH; the dominant species at physiological pH is the dimeric one. This species is highly stable, as shown by thermal denaturations followed by DSC, CD, and fluorescence. The GdmCl chemical denaturations at physiological pHs are reversible and exhibited two-state behavior. Then, the mechanism by which alternative quaternary structure arises does not involve the formation of fluctuating monomeric species, but rather docking on a preformed oligomeric species.

EXPERIMENTAL PROCEDURES

Materials. Ultrapure GdmCl was from ICN Biochemicals. Exact concentrations of GdmCl and urea were calculated from the refractive index of the solution (*16*). Standard suppliers were used for all other chemicals. Water was deionized and purified on a Millipore system.

Protein Purification. Protein was purified as described in ref *15*. Protein concentrations were calculated from the absorbance of stock solutions measured at 280 nm, using the extinction coefficients of model compounds (*17*). Protein purities were confirmed by MALDI-TOF. Briefly, samples were mixed with a saturated solution of α-cyano-4-hydroxy-cinnamic acid in a 2:1 (v/v) ratio. One microliter of the mixture was spotted in each well of the stainless steel target plate and cocrystallized by evaporation. Spectra were acquired in positive linear mode with the aid of a Bruker Reflex IV time-of-flight mass spectrometer.

Fluorescence Measurements. Fluorescence spectra for MJ0729 were recorded in a Cary Eclipse spectrofluorometer (Varian) interfaced with a Peltier cell. The sample concentration was in the range of 1−3 μM, and the final concentration of the buffer was, in all cases, 10 mM. A 1 cm path length quartz cell (Hellma) was used.

(a) Steady State Measurements. Samples of MJ0729 were excited at 280 and 295 nm at 25 °C in the pH range of 3.0−12.0 to characterize a putative different behavior of tryptophan or tyrosine residues (*17*). The slit width was typically equal to 5 nm for the excitation and 5 nm for the emission light. The fluorescence experiments were conducted between 300 and 400 nm. The signal was acquired for 1 s, and the wavelength increment was 1 nm. Blank corrections were made in all spectra.

(b) Chemical and pH Denaturation Experiments. In the GdmCl denaturation experiments, fluorescence spectra were acquired with the parameters described above. Spectra were corrected by subtracting the baseline in all cases. The chemical denaturation reaction was fully reversible at pH 7.0, as shown by the sigmoidal curves obtained by starting from diluted 7 M GdmCl samples at different pH values (data not shown). All the samples were left overnight to equilibrate at 4 °C. Every chemical denaturation experiment was repeated at least three times with new samples. Chemical denaturations were irreversible below pH 5.0 (data not shown). The protein concentration was 1.26 μM, in protomer units, unless it is stated (in the following, we will identify protomer as a monomer of the MJ0729 protein). The following salts and acids were used: phosphoric acid at pH 2.0−3.0, formic acid at pH 3.0−4.0, acetic acid at pH 4.0−5.5, NaH_2PO_4 at pH 6.0−7.0, Tris acid at pH 7.5−9.0, Na_2CO_3 at pH 9.5−11.0, and Na_3PO_4 at pH 11.5−13.0. The pH was measured with an ultrathin Aldrich electrode in a Radiometer (Copenhagen, Denmark) pH-meter.

Circular Dichroism Spectroscopy. Circular dichroism spectra were recorded on a Jasco J810 spectropolarimeter fitted with a thermostated cell holder and interfaced with a Neslab RTE-111 water bath. The instrument was periodically calibrated with (+)-10-camphorsulfonic acid.

(a) Steady State Measurements. Isothermal wavelength spectra at different pHs were acquired at a scan speed of 50 nm/min with a response time of 2 s and averaged over four scans at 25 °C. Far-UV measurements were performed with

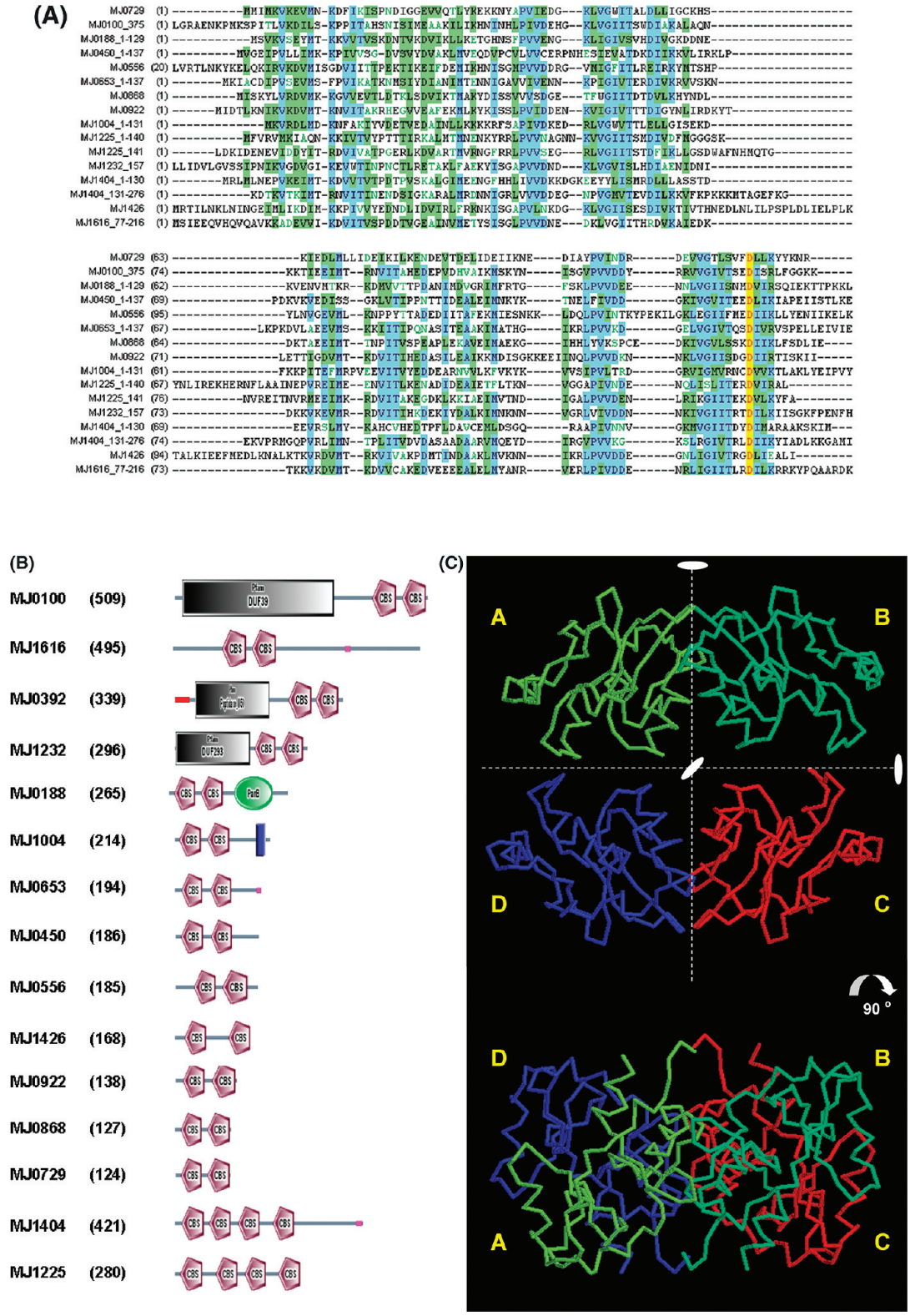

FIGURE 1: Sequence comparison and distribution of CBS domains and CBS domain-containing proteins in *M. jannaschii*. (A) Sequence alignment of the CBS domain proteins. The figure shows a CLUSTALW alignment (Align X from Vector NTI 9.0, InforMax) of exclusively the CBS domain pair fragment of the proteins identified in the *M. jannaschii* genome as carriers of this conserved domain. The number within parentheses for each indicates the number of the first amino acid in the alignment. Color codes are as follows: red on yellow background, invariant amino acids; black on blue, strongly conserved; black on green, similar; green on white, weakly similar; black on white, not conserved. The protein sequences shown have the following accession numbers: MJ0729 (NP_247714), MJ0100 (NP_247064), MJ0188 (NP_247156), MJ0450 (NP_247424), MJ0566 (NP_247535), MJ0653 (NP_247637), MJ0868 (NP_247863), MJ0922 (NP_247917), MJ1004 (NP_247998), MJ1225 (NP_248220), MJ1232 (NP_248227), MJ1404 (NP_248407), MJ1426 (NP_248430), and MJ1616 (NP_248626). (B) Motif distribution in CBS domain-containing proteins (SMART database, http://smart.embl-heidelberg.de/). Chain lengths are given in parentheses. (C) Cα atom representation of tetrameric MJ0729 as determined from the crystal structure of the protein at pH 4.5−4.8. The 2-fold pseudosymmetry axis relating the four monomers is indicated with a dotted line. A putative dimer of the protein would be represented by subunits A and D or B and C. Panel C was created with RASMOL (http://www.umass.edu/microbio/rasmol/). The bottom part of the panel depicts a rotation of 90° around the horizontal axis to show the oblong shape of tetrameric MJ0729.

10−15 μM protein in 10 mM buffer, using 0.1 cm path length quartz cells (Hellma). Near-UV spectra were acquired with 30−40 μM protein in a 0.5 cm path length cell. All spectra were corrected by subtracting the proper baseline. The mean residue ellipticity, [θ], was calculated as described previously (*18*). The helical content of MJ0729 at any pH was approximated from its mean residue ellipticity at 222 nm as described previously (*18*).

(b) Thermal Denaturation. Thermal denaturation experiments with MJ0729 were performed at constant heating rates of 60 °C/h and a response time of 8 s. They were collected in the far-UV region by following the ellipticity at 222 nm from 25 to 95 °C in 0.1 cm path length quartz cells (Hellma) with a total protein concentration of 20−40 μM. The reversibility was tested by recording a new scan after the thermally denatured samples had cooled to 25 °C and comparing it with the spectra obtained in the first scan. In all cases, both spectra were identical (data not shown). The possibility of drifting of the CD spectropolarimeter was tested by running two samples containing only buffer, before and after the thermal experiments. No difference between the scans was observed. Every thermal denaturation experiment was repeated at least twice with new samples. In the reheating experiments, the samples were always transparent and no precipitation was observed.

In the pH-induced unfolding experiments, the pH was measured after completion of the experiments, and essentially no differences were observed with those pHs calculated from the buffer stock solutions. The pH range explored was 3.0−12.0. The proper blank solutions were subtracted in all cases. The buffer concentration was 10 mM in all cases, and the buffers were the same used in the fluorescence measurements. Samples were left overnight to equilibrate at 4 °C.

Differential Scanning Calorimetry. DSC experiments were carried out with a VP-DSC calorimeter (MicroCal, Northampton, MA). Protein solutions were prepared by exhaustive dialysis against the working buffer [50 mM HEPES, 200 mM NaCl, and 2 mM β-ME (pH 7.0)] at 4 °C. To minimize the amount of gas dissolved in the solutions, all the samples were degassed under vacuum for 10 min at room temperature with gentle stirring before being loaded into the calorimetric cells. Samples were heated at a constant scan rate of 1.5 °C/min (90 °C/h) and held under an extra pressure of 28 psi to avoid bubble formation and evaporation at temperatures above 95 °C. Several buffer−buffer scans were performed to ensure proper instrument equilibration. To test whether the heat-induced denaturation of the protein was reversible, protein solutions were cooled in situ to 20 °C for 30 min immediately after the first scan was completed (usually ranging from 20 to 125 °C) and rescanned under the same experimental conditions. To correct for small mismatches between the two cells, an instrumental baseline (i.e., buffer−buffer baseline) was subtracted from the protein endotherm before data analysis. All traces were dynamically corrected to account for the time-delayed response of the detector to the heat event that evolved within the calorimetric cell. After normalization to concentration, a chemical baseline calculated from the progress of the unfolding transition was subtracted.

Fitting was carried out by using the Origin 7.0 package supplied with the instrument. Since we found no evidence of dissociation coupled to the thermal denaturation of the protein (single symmetrical and concentration-independent endotherm; see Results), the excess heat capacity functions were fitted to the two-state (nondissociating) model. The unfolding of the protein was shown to be complicated by its precipitation at temperatures above its T_m, especially at protein concentrations higher than 0.5 mg/mL. In those cases, since the contribution of the irreversible process leads to the distortion of the endotherm and therefore deviations from the two-state model, no fitting of the data was attempted and only the total calorimetric enthalpy and the maximum in the excess heat capacity were considered relevant.

Fourier Transform Infrared Spectroscopy. Infrared spectra were recorded in triplicate (600 spectral scans) in a Bruker IF66s instrument equipped with a DTGS detector.

(a) Steady State Measurements. For infrared amide I′ band recordings, aliquots of 400 μg of protein in 10 mM HEPES were washed twice with 2 mL of 10 mM MOPS, 10 mM CAPS, 10 mM potassium acetate (pH 7.0), 100 mM KCl or 10 mM MOPS, 10 mM CAPS, 10 mM potassium acetate (pH 2.5), and 100 mM KCl, and their volumes were reduced to ~20 μL by filtration on Vivaspin concentrators [5000 MW cutoff (Vivascience)]. The concentrated samples were dehydrated in a speedvac Savant rotary evaporator and resuspended in 20 μL of D_2O to prevent the interference of H_2O infrared absorbance (1645 cm^{-1}). The resulting samples were placed in a liquid demountable cell equipped with CaF_2 windows and 50 μm thick Mylar spacers and maintained at room temperature for approximately 2 h to reach equilibrium. The buffer contribution was subtracted from the individual spectra, and spectral noise was reduced as described previously (*19*).

(b) Thermal Denaturation Measurements. The samples were subjected to heating cycles at the indicated temperatures. Each step includes (i) a steplike increase in temperature, (ii) a stabilization period of the sample (or plain buffer) in the IR cell at each temperature, and (iii) a period of spectral acquisition. The duration of a complete heating cycle was approximately 2.5 h.

(c) Band Fitting Analysis. The protein secondary structure was estimated from the IR spectra by decomposition of the amide I′ band into its spectral components (*20*). Spectral smoothing was carried out by applying the maximum entropy method, assuming that noise and band shape follow a normal distribution. The minimum bandwidth was set to 12 cm^{-1}. The signal-to-noise ratio of the processed spectra was better than 10000:1. Derivation of IR spectra was performed using a power of 3, with a breakpoint of 0.3; Fourier self-deconvolution was performed using a Lorenztian bandwidth of 18 cm^{-1} and a resolution enhancement factor of 2.0.

Nuclear Magnetic Resonance Spectroscopy. ^1H NMR experiments were conducted in a Bruker Avance DRX-500 instrument.

(a) One-Dimensional NMR Spectra. Spectra were recorded with 32K data points and using the WATERGATE sequence to eliminate the residual water signal (*21*). Typically, 1024 scans were acquired, and the spectral width was 8000 Hz in all cases. Samples in D_2O were prepared with Amicon centrifugal devices in 10 mM phosphate buffer (pH 7.5). Spectra were acquired at different concentrations ranging from 0.1 to 1.5 mM at 25 °C. No changes in chemical shifts or line broadening were observed. The spectra were processed by using BRUKER-UXNMR working on a Windows

workstation. An exponential window function and polynomial base line corrections were applied. The final one-dimensional spectra contained 64000 data points. ^{1}H chemical shifts were quoted relative to external TSP.

(b) Translational Diffusion Measurements (DOSY experiments). Translational self-diffusion measurements were performed using the pulsed-gradient spin-echo NMR method (*22*). The following relationship exists between the translational self-diffusion parameter, *D*, and the NMR parameters (*23*):

$$\frac{I}{I_0} = -\exp\left[D\gamma_H{}^2\delta^2G^2\left(\Delta - \frac{\delta}{3} - \frac{\tau}{2}\right)\right] \quad (1)$$

where I is the measured peak intensity (or volume) of a particular (or a group of) resonance(s), I_0 is the maximum peak intensity of the same (group of) resonance(s) at the smaller gradient strength, D is the translational self-diffusion constant (in square centimeters per second), δ is the duration (in seconds) of the gradient, G is the strength of the gradient (in teslas per centimeter), Δ is the time (in seconds) between the two gradients (i.e., the time when the molecule evolves), and τ is the time (100 μs) for gradient recovery at the end of each scan. Data can be plotted as the $-\ln(I/I_0)$ versus G^2, and the slope of the line is $\gamma_H{}^2\delta^2G^2(\Delta - \delta/3 - \tau/2)$, and then D can be easily obtained.

The Stokes−Einstein equation relates D to the molecular shape via the so-called friction coefficient, f:

$$D = kT/f \quad (2)$$

where T is the temperature (in kelvin) and k Boltzmann's constant. The f of a protein is determined by its overall dimensions, hydration, and the rugosity of the surface exposed to water. If it is assumed that the protein adopts a spherical shape, the f coefficient is given by

$$f = 6\pi\eta R_s \quad (3)$$

where η is the viscosity of the solvent and R_s is the hydrodynamic radius of the sphere. From eqs 2 and 3:

$$R_s = kT/6\pi\eta D \quad (4)$$

The viscosity of a solution is very weakly influenced by the macromolecule component at the low macromolecular concentrations used, and therefore, the viscosity of the solution should be that of the solvent. Solvent viscosity is temperature-dependent according to the relationship (*24*) $\log \eta = a + (b/c - T)$. The terms a, b, and c are given for a particular D_2O:H_2O ratio. Under our conditions, a 100% D_2O solution, at 20 °C, the values of a, b, and c are −4.2911, −164.97, and 174.24, respectively. This yields an η value of 1.253 kg cm^{-1} s^{-1} at 20 °C, used in our calculations.

The gradient strength was calibrated by using the diffusion rate for the residual proton water line in a sample containing 100% D_2O in a 5 mm tube, and back-calculating G. This procedure assumes that the diffusion rate for HDO in a 100% D_2O sample is 1.94×10^{-5} cm^2 s^{-1} at 25 °C (*24*). Experiments were conducted by using the longitudinal eddy current delay pulse sequence, with a postgradient eddy current relaxation delay of 5 ms. Each experiment was averaged over 128 scans, and the number of points was 16K. The strength of the gradients was varied from 2% of the total power of the gradient coil to 95%, and they were sine-wave-shaped. Experiments were conducted at different protein concentrations; the largest protein concentration used was 1 mM. The other concentrations were obtained from dilution of the 1 mM stock. The duration of the gradient was varied between 3 and 2.2 ms, and the time between both gradients was changed between 100 and 150 ms. The most upfield-shifted methyl groups (between 0.5 and 0.8 ppm) were used to measure the changes in intensity. Then, since the chemical shift is an averaged parameter through all the species present in solution, the determined D is the average for all species present in the solution at any pH.

Gel Filtration Chromatography. Experiments were carried out in an analytical gel filtration Superdex 75 HR 16/60 column (GE Healthcare) running on an AKTA FPLC system at 25 °C and observed at 280 nm. Flow rates of 1 mL/min were used, and aliquots of 100 μL were loaded into the column after equilibration. The column was equilibrated with 4 column volumes of elution buffer. The elution buffer contained 150 mM NaCl added to avoid nonspecific interactions with the column and the corresponding buffer (at 50 mM). No differences in the elution volumes were observed among the different protein concentrations used (from 4 to 100 μM). The column was calibrated using the gel filtration low-molecular mass calibration kit (GE Healthcare). The following protein standards with their corresponding Stokes radii were used: ribonuclease A (1.64 nm), chymotrypsinogen (2.09 nm), ovoalbumin (3.05 nm), and bovine serum albumin (3.55 nm) (*25*). Each experiment at the different pHs was repeated three times with fresh new samples.

The elution of a macromolecule in gel filtration experiments is usually given by the partition coefficient, which is defined as the fraction of solvent volume within the gel matrix accessible to the macromolecule (*18*). The weight average partition coefficients (σ) of protein standards and MJ0729 were calculated by

$$\sigma = \frac{(V_e - V_o)}{V_i} \quad (5)$$

where V_e is the elution volume of the protein and V_o and V_i are the void and internal volumes of the column, with values of 8.13 ± 0.06 and 28.43 ± 0.03 mL, respectively. V_o and V_i were determined using riboflavin (5 mg/mL, in 10 mM phosphate buffer containing 150 mM NaCl) and L-tryptophan (0.5 mg/mL, in the same buffer), respectively, by averaging four measurements.

The partition coefficients were determined for the molecular size standards, and they were transformed by using the inverse error function complement of σ, $\mathrm{erfc}^{-1}(\sigma)$, to yield a linear relationship with the molecular Stokes radius, R_s (*18*):

$$R_s = a + b[\mathrm{erfc}^{-1}(\sigma)] \quad (6)$$

where a and b are the calibration constants for the column.

Fitting of the calculated $\mathrm{erfc}^{-1}(\sigma)$ to eq 6 by linear least-squares analysis was carried out on Kaleidagraph (Abelbeck Software) on a personal computer. Once the calibration parameters are obtained, the Stokes radius of any macromolecule can be determined.

For a protein, r_o, the radius of the spherical anhydrous macromolecule, can be calculated considering that the

anhydrous molecular volume ($M\bar{V}/N$) equals the volume of a sphere (*26*):

$$\frac{M\bar{V}}{N} = \frac{4}{3}\pi r_o^3, \text{ which yields } r_o = \sqrt[3]{\left(\frac{3M\bar{V}}{4N\pi}\right)} \qquad (7)$$

where M is the molecular mass of the protein, $\bar{V}$ is the partial specific volume of the protein, and N is Avogadro's number. The $\bar{V}$ of MJ0729 is 0.73 cm^3/g, and its molecular mass is 14302.7 Da.

Dynamic Light Scattering. The oligomerization state and the homogeneity of the purified protein were analyzed by using dynamic light scattering (DLS). All experiments were carried out in Dynamic Light Scattering DLS Dynapro Titan equipment (Wyatt-Tecnology) at 20 °C at a laser wavelength of 620 nm. The protein was divided in eight samples and equilibrated at different pHs by using "Protein Desalting Spin Columns" (Pierce). Protein concentrations were measured with a NanoDrop at 280 nm. Experiments were carried out at pHs between 3.0 and 11.0, using protein concentrations from 10 to 120 mg/mL. The following buffers were used: 100 mM citric acid for the pH range of 3.0−6.0, 100 mM Hepes for the pH range of 7.0−8.0, and 100 mM glycine for the pH range of 9.0−11.0. Samples were filtered by using a 0.1 mm filter (Whatman, Maidstone, U.K.). The correlation function was measured 10 times, over 10 s, and averaged to yield the result of the hydrodynamic radii. Each sample was measured three times.

DLS measures time-dependent fluctuations in the intensity of scattered light, which occurs because particles are undergoing random Brownian motion in the solution. The intensity of scattered light, I, is proportional to the sample concentration, C, and the molecular mass of the molecules, M, in the solution, according to

$$I \approx MC\left(\frac{\delta n}{\delta C}\right)$$

where n is the refraction index of the solution.

In DLS experiments, the intensity autocorrelation function of the scattered light, $g^2(\tau)$, is measured according to a delay time, τ, following the relationship $g^2(\tau) = [\langle I(t)I(t+\tau)\rangle]/[\langle I(t)\rangle^2]$. This function is then used to calculate the normalized time correlation function of the scattered electric field (*27*), $g^1(\tau)$. In practice, τ is related to a fundamental sample time, Δt, to perform the dynamic study of the signal. The function $g^1(\tau)$ in monodisperse particles can be analyzed with a Cummulant expansion method (*28*) following a single exponential: $g^1(\tau) = b + a\exp(-\Gamma\tau)$, where Γ is the decay rate function, which is obtained with the software of the instrument (Dynamics 6.7.7.9). This software uses a nonlinear least-squares fitting algorithm to fit the measured correlation function. The decay rate is related to the D with the relationship $\Gamma = Dq^2$, where q is the scattering vector and D is the diffusion coefficient defined above (*29*), from which the R_s can be obtained (eq 4). However, it is important to indicate that the measured D obtained by this procedure is the average for all species present in the solution at any pH.

In all samples analyzed, several peaks, even with high molecular radii, were observed, but only the most intense one was used for analysis (see Figure 2A). All the samples used at pH 3.0, 5.0, and 7.0 at any of the protein concentrations explored were monodisperse (<20% polydispersity); at pH 6.0 and 8.0, the most concentrated samples exhibited between 20 and 30% polydispersity, but in all of the samples, the polydispersity was larger than 40%, which is the limit to admit polydispersity (*28*).

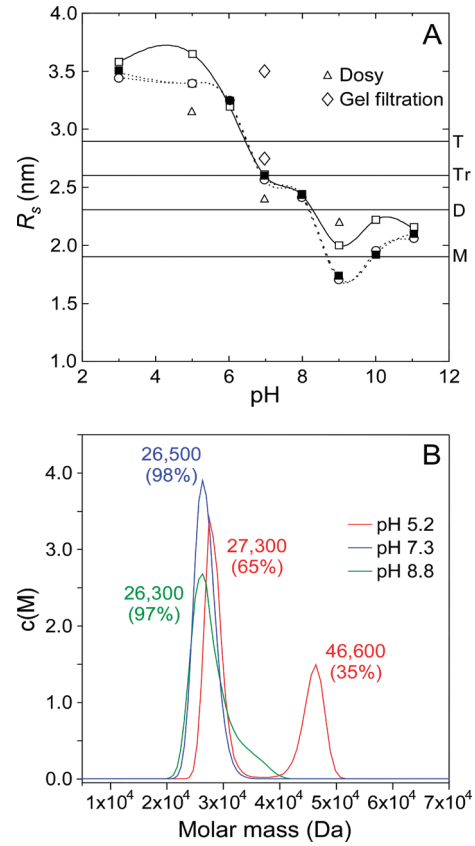

FIGURE 2: Hydrodynamic properties of MJ0729. (A) Variation of the hydrodynamic radius (R_s) of MJ0729 with pH and protein concentration [(□) 80, (■) 50, and (○) 10 mg/mL], as measured by DLS at 20 °C. The curve is drawn to guide the eye. The straight horizontal lines correspond to the theoretical R_s values corresponding to monomer (M), dimer (D), trimer (Tr), and tetramer (T), with one-shell (0.3 nm) hydration water; the values are 1.6 ± 0.3, 2.0 ± 0.3, 2.3 ± 0.3, and 2.6 ± 0.3 nm, respectively (eq 7). Estimated errors in the DLS measurements are 10%. To allow for comparison, the results from DOSY-NMR (△) (Table 1) and gel filtration experiments (◇) (Figure 3B) are shown. (B) Concentration distribution vs molecular mass (daltons) of the association states of MJ0729 at pH 5.2, 7.3, and 8.8 and 20 °C. The percentages of the different species are indicated.

Table 1: Infinite-Dilution Diffusion Coefficients of MJ0729[a]

pH	D ($\times 10^6$ cm^2 s^{-1})	R_s (nm)[b]
5.0	5.5 ± 0.8	3.2
6.8	8.0 ± 0.6	2.4
8.8	8.3 ± 0.5	2.2

[a] Errors are fitting errors from the linear extrapolation at infinite dilution. Measurements were taken at 20 °C in D_2O. [b] The hydrodynamic radius was obtained from eq 4.

Analytical Ultracentrifugation. The sedimentation velocity experiments were conducted in an XL-A analytical ultracentrifuge (Beckman-Coulter Inc.) at 42000 rpm and 20 °C, using an A50Ti rotor and a 3 mm charcoal-filled Epon double-sector centerpiece. Absorbance was measured at 280 nm. The protein concentration was 20 μM in 50 mM Tris-HCl (pH 7.3). Data were modeled as a superposition of Lamm equation solutions with SEDFIT (available at ww-

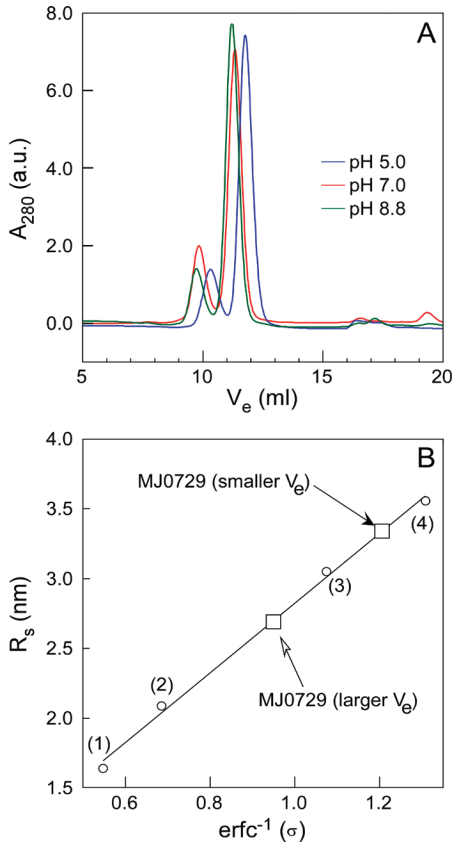

FIGURE 3: Gel filtration measurements of MJ0729. (A) Elution peaks of MJ0729 at different pHs and protein concentrations. Experiments at pH 5.0 were conducted at a protomer protein concentration of 100 μM [similar results were obtained at 4 μM protomer (data not shown)]; experiments at pH 7.0 and 8.8 were conducted at 4 μM protomer [similar results were obtained at 100 μM protomer (data not shown)]. (B) Determination of the R_s of MJ0729 at pH 7.0. The shortest elution volume peak is indicated by a filled arrow and that appearing at largest volumes by a blank arrow. The numbering corresponds to the elution volumes of ribonuclease A (1), chymotrypsinogen (2), ovalbumin (3), and bovine serum albumin (4). The equation was $R_s = (2.5 \pm 0.1) \times \text{erfc}^{-1}(\sigma) + (0.3 \pm 0.1)$ (with a regression coefficient of 0.99). Experiments were conducted at 25 °C.

w.analyticalultracentrifugation.com/default.htm) (*30*). The sedimentation coefficient distribution, $c(s)$, was calculated at a $p = 0.68$ confidence level. The experimental sedimentation values were determined by integration of the main peak of $c(s)$ and corrected to standard conditions to obtain the $s_{20,w}$ values with SEDNTERP (*31*). Calculation of the frictional coefficient ratio was performed with SEDFIT to yield the $c(M)$ distribution (*30*). Experiments were also conducted at pH 5.2 (50 mM acetic acid) and pH 8.8 (50 mM Tris).

Analysis of the pH, Chemical, and Thermal Denaturation Curves, and Free Energy Determination. The average emission intensity, $\langle\lambda\rangle$, in fluorescence spectra was calculated as (*32*) $\sum_1^n[(1/\lambda_i)/I_i]/\sum_1^n I_i$, where I_i is the fluorescence intensity measured at wavelength λ_i. The pH denaturation experiments were analyzed assuming that both species, protonated and deprotonated, contributed to the spectrum:

$$X = \frac{X_a + X_b \times 10^{n(\text{pH}-\text{p}K_a)}}{1 + 10^{n(\text{pH}-\text{p}K_a)}} \quad (8)$$

where X is the physical property observed (ellipticity, fluorescence intensity, or $\langle\lambda\rangle$), X_a is that in the acidic region, X_b is that at high pHs, pK_a is the apparent midpoint of the titrating group, and n is the Hill coefficient. The apparent pK_a was obtained from the average of three different measurements. Fitting to eq 8 by nonlinear least-squares analysis was carried out by using Kaleidagraph (Abelbeck Software) on a personal computer. Global fittings were carried out with MATLAB (Prentice-Hall).

The GdmCl denaturation curves were analyzed as described in other oligomeric proteins (*33*), calculating an apparent free energy as described by Marqusee and co-workers (*34*). During the fitting of chemical denturation data, we used the two-state model with a protein concentration-dependent parameter as described by Bönisch and co-workers (*33*). Briefly, these authors have proposed that the term $-RT \ln[(n^n/2^{n-1})/c_t^{n-1}]$, where n is the oligomerization state of the protein (in MJ0729, it should be 2 or 4; see Discussion) and c_t is the molar concentration of the protein expressed in n-mer equivalent; this term is included in the expression given for the apparent free energy as described by Marqusee (*34*). We did not use a three-state model, as those described by other authors (*35*), since we did not find evidence of an intermediate species (that is, in MJ0729, the chemical denaturation midpoints determined by CD and fluorescence agree within the experimental uncertainty), and it could result in an overinterpretation of the results. Thermal denaturation data (at the pHs where a T_m could be obtained; see below) were analyzed using a concentration-dependent two-state unfolding mechanism (*33*).

RESULTS

The Oligomerization State of MJ0729 Is pH-Dependent. We have previously described the overexpression, crystallization, and X-ray preliminary analysis of protein MJ0729 (*15*). The protein yielded complex twinned trigonal crystals at low pHs, which ranged from hemihedrally to tetartohedrally twinned species. X-ray grade crystals were often difficult to obtain and diffracted poorly to low or medium resolution. During the crystallization, we observed that the shape and quality of the crystals varied significantly with slight variations in the pH of the buffer. We have recently determined the crystal structure of MJ0729 (which is being reported elsewhere) which strikingly adopts a disklike shape tetrameric fold with dimensions of 65 Å × 65 Å × 38 Å (Figure 1C). This fold and dimensions are different from those found in related proteins and affect the shape of the ligand binding site cavity and potentially its ligand binding capabilities, probably due to the nonconventional arrangement adopted by the Bateman domains (Figure 1C).

Briefly, MJ0729 is an α/β structure that consists of two CBS domains associated in tandem generating a simple compact structure known as a "Bateman domain" that was first described in inosine monophosphate dehydrogenase (IMPDH) (*1, 2*). Accordingly, each CBS domain of MJ0729 is formed by a three-stranded antiparallel β sheet and two α helices packed against them, with a β1-α1-β2-β3-α2 sequence (*2*). In the crystal, MJ0729 forms a compact homotetramer with an extensive interface and internal pseudosymmetry 222 (Figure 1C), which favors the tetartohedral twinning of the crystal (*15*). The four subunits constituting one tetramer have been named A, B, C, and D (Figure 1C), in which A and B (or C and D), A and C (or B and D), and

A and D (or B and C) are related by a pseudo-2-fold axis parallel to the *a*, *b*, and *c* axes of the unit cell, respectively (Figure 1C). Thus, the tetrameric structure of MJ0729 could be described as the association of two dimers (i.e., monomers A−D + monomers B−C in Figure 1C) related by a pseudo-2-fold axis (represented with a vertical dotted line in Figure 1C). No contacts occur between subunits A and C (or B and D). An alternative interpretation for a potential dimer would involve the association of subunits A and C or B and D. Thus, taking into account the current knowledge from related structures, the association of subunits A and D (or B and C) seems more feasible as a potential dimer of MJ0729. Strikingly, and in contrast with other stand-alone CBS domain proteins from prokaryotic and eukaryotic organisms, in tetrameric MJ0729 the A−D Bateman pairs (and/or B−C) are arranged in a manner remarkably different from the standard head-to-head or head-to-tail dimeric associations in which two Bateman domains lay within the same plane. These structures are exemplified by Protein Data Bank (PDB) entries 1PBJ and 1050 (head-to-tail dimers) and 1PVM and 1YAV (head-to-tail dimers). A common feature of these dimeric CBS tandem pairs is the hydrophobic surface of the αA and αB helices, which mediates their self-association (*6*). Our crystallographic analyses of MJ0729 have revealed a novel tetrameric association in which the Bateman pairs configuring each potential dimer (A−D or B−C in Figure 1C) are contained within two perpendicular planes. In addition, within each dimer, the Bateman2 domain is tilted with respect to the Bateman1 domain, disrupting the interaction among αA and αB helices usually observed in standard dimers. Then, MJ0729 would represent an intermediate situation between a head-to-head and head-to-tail arrangement for each A−D and/or B−C dimer. As a result, an unusual ligand binding cavity is formed in which residues from a major loop of the second Bateman domain instead of those from helices αA or αB are close to the lower part of the crevice. According to the crystal structure, one might speculate that the MJ0729 tetramer would be stabilized upon formation of new specific interactions between amino acid residues of A−B monomers (or C−D) from the corresponding dimers; the presence of these new interactions might modify the former orientation of the Bateman domains within each dimer, which could formerly be organized according to a standard head-to-head or head-to-tail arrangement.

Then, to elucidate whether the observed tetrameric species is a crystallization artifact, we carried out hydrodynamic measurements of MJ0729 at different pHs. In turn, this study has been crucial for improving the quality of the crystals.

DLS Measurements. Theoretically, the radius for a monomeric spherical MJ0729 (14302.7 Da) was 1.6 nm (eq 7), that for dimeric MJ0729 2.0 nm, that for a trimeric species 2.3 nm, and that of a tetrameric one 2.6 nm. However, to allow comparison with the radius determined by the hydrodynamic techniques in this work, we should add to the calculated values a correction of 0.3 nm to account for the hydration shell (*26*) (Figure 2A). Comparison of the experimental R_s values with the theoretical ones suggests that (i) at physiological and basic pHs there is probably an equilibrium between a dimeric species and another oligomer; and (ii) at low pH there is an equilibrium involving long-R_s species. Since the tetramer at pH 4.5−4.8 obtained by X-ray is an oblong molecule (Figure 1C), we favor that some of those long-R_s species at low pH are tetramers (see Discussion), but not all (as suggested by the FTIR measurements; see below). It is important to indicate that at basic pH, we cannot be sure of the existence of a monomeric species, since the measured *D* is the result of all the species present in solution, whatever the shape and oligomerization state. It could be that the smaller measured *D* at basic pHs is due to the presence of a more compact dimeric species, which is formed upon rearrangement of the physiological dimeric ones. No significant differences in the R_s values from those from the DLS measurements were observed at the three protein concentrations that were used (Figure 2A).

DOSY-NMR Measurements. The pHs explored by NMR measurements were 5.0, 6.8, and 8.8. The *D* at infinite dilution was pH-dependent, with the largest values at basic pHs (Table 1). At all pHs, we observed a straight line in the measured *D* coefficients, suggesting that no concentration dependence was being observed in this concentration range (from 5.4 to 0.5 mg/mL) (data not shown). The hydrodynamic radii obtained from eq 4 at any pH were similar to those calculated by DLS (Figure 2A). Then, the deviations observed in both techniques from the ideal spherical radii are probably due to the oblong nature of the molecules involved or equilibria involving different species. Then, the presence of any possible monomeric species at basic pHs cannot be ensured by this technique.

Gel Filtration Experiments and Shape Parameters. MJ0729 eluted at any of the explored pHs as two peaks (Figure 3A) at either 4 or 100 μM. At physiological and basic pH, these peaks appeared at 9.8 (the smaller one) and 11.36 mL (the largest one), which resulted in R_s values of 3.4 and 2.6 nm, respectively (Figure 3B). The first radius is larger than that theoretically calculated for a tetrameric sphere, and the second one was close to that of a trimeric sphere. However, it must be kept in mind that gel filtration chromatography separates molecules not only by molecular weight but also by shape. These results suggest that notwithstanding the presence of the matrix, the behavior of the protein at these pHs is similar to those obtained by the other techniques.

On the other hand, the species eluting at pH 5.0 exhibited larger V_e values and smaller hydrodynamic radii (Figure 3A). At acidic pHs, DLS and DOSY results show the presence of long-R_s species, which in a gel filtration experiment should elute at smaller volumes. Then, the findings in the gel filtration at pH 5.0 can be explained as being only due to protein−column interactions, probably because the protein is not completely folded at this pH and has solvent-exposed hydrophobic patches, as indicated by the intrinsic fluorescence and ANS binding experiments (see below).

Analytical Ultracentrifugation. The results from the hydrodynamic techniques described above suggest the presence of long-R_s species at acidic pHs, which are probably self-associated ones, with solvent-exposed hydrophobic patches. However, the hydrodynamic techniques used so far cannot distinguish contributions of mass and shape to molecular diffusion. Conversely, analytical ultracentrifugation can be used to determine directly the molar mass of the association state of macromolecules (*36*). The oligomeric state of MJ0729 was investigated with a sedimentation equilibrium at the same pHs where DLS, gel filtration, and DOSY measurements were conducted.

In the sedimentation velocity experiment at pH 8.8, a single species was observed with an estimated molecular mass of 26.3 kDa (Figure 2B), similar to that obtained for a dimeric species (28.6 kDa); however, a tailing in the $c(M)$ was observed at these high pHs (green line in Figure 2B), suggesting that probably there is an equilibrium between dimeric species with different levels of compactness (which reinforces our previous hypothesis discussed above in the DLS results and supports the idea that no monomeric species are present at these basic pHs).

A similar value of molecular mass was obtained at pH 7.3 (26.5 kDa). However, at pH 5.2, two populations were observed: that corresponding to a dimer (27.3 kDa) and another (46.6 kDa), which is halfway between the molecular masses of trimeric and tetrameric species (the trimeric species should have a theoretical molecular mass of 42.9 kDa).

The sedimentation coefficient of a macromolecule, s, and the molecular parameters that determine the s value are given by the well-known Svedberg equation:

$$s = \frac{MD(1 - \bar{V}\rho)}{RT}$$

where ρ is the density of the solvent. Experimentally, we obtain the s parameter from the sedimentation velocity experiments, but we cannot determine experimentally the value of D. On the basis of a deconvolution of the diffusion effects, a differential sedimentation coefficient distribution, $c(s)$, can be defined as

$$a(r, t) = \int c(s)\chi[s, D(s), r, t]\, \mathrm{d}s$$

where $a(r,t)$ is the observed sedimentation data, $c(s)$ is the concentration of species with a sedimentation coefficient between s and $s + \mathrm{d}s$, and $\chi[s,D(s),r,t]$ is the solution to the underlying transport equation (Lamm equation) (*36*), which is also a function of time and radial position. SEDFIT solves the equation given above by using maximum entropy methods, with several approaches to estimate the relationship of s and D, and then to obtain from the $c(s)$, the molar distribution $c(M)$. The first approach decomposes D as a monotonous single-value function of s. The second, more common, assumption is that all sedimenting species have the same weight average frictional shape factor [the rate between the f factor (eq 3) and that of a spherical unsolvated molecule] (*37*). This latter approach gives estimated molecular masses within 10% of that of the target protein, but caution must be kept in mind since $c(M)$ will be always dependent on that frictional shape factor. We think that this is the reason why the calculated molecular mass for that other species present at acidic pH is halfway between that of a trimer and that of a tetramer (Figure 2B). Due to the titration of some acidic groups, the shape of the tetrameric species must be highly distorted (or alternatively the shape of the resulting dimeric species present in solution must be highly elongated), resulting in different oligomeric species whose frictional shape factor must be quite different among them and with the rest of the species in solution. Then, the apparent trimeric species at acidic pHs is a tetrameric one, as suggested (but not unambiguously shown) by the other hydrodynamic techniques. At this stage, however, we cannot rule out the presence of a trimeric species at these pHs, but we find it rather difficult to rationalize the dissociation of a tetrameric species to a trimeric one, and from this to a dimer protein, though it is interesting to note that such behavior has been observed in other proteins (see Discussion).

Then, to sum up, the hydrodynamic techniques indicate that the quaternary structure of MJ0729 is pH-dependent, with a dominant population of dimeric species at physiological pHs, and a long-R_s species at slightly acidic pHs (pH 4.5).

The Structure of MJ0729 Remained Unaltered over a Wide pH Range. The results indicate that the quaternary structure was pH-dependent, but what happens with the secondary and tertiary structures? To address this question, we carried out a spectroscopic characterization at different pHs.

Intrinsic Fluorescence. We used the intrinsic fluorescence spectra to map any change in the tertiary structure of the protein with pH. The monomer of MJ0729 has a sole tryptophan at position 51 in the first CBS domain and five tyrosines (two in the first CBS domain and three in the second). The emission fluorescence spectrum of MJ0729 at pH 7.0, by excitation at 280 or 295 nm, was blue-shifted, with a maximum at 326 nm. This indicates that the tryptophans were completely buried in the structure and had the largest contribution to the spectrum. At pH <5.0, the maximum of the spectrum was red-shifted toward 338 nm; a similar behavior was observed at pH >9, suggesting that at the extremes of pH the tryptophans became solvent-exposed. A similar tendency was also observed in $\langle\lambda\rangle$ (Figure 4A, ■, right axis) and in the intensity at 326 nm (Figure 4A, □, left axis). These findings suggest that at the extremes of pHs, the tryptophans and probably some, if not all, of the tyrosine residues became solvent-exposed. The pK_a of the acidic tritration determined by the changes in intensity at 326 nm was 4.9 ± 0.2.

ANS Fluorescence. ANS binding was used to monitor the solvent-exposed hydrophobic regions (*38*). At low pHs, the fluorescence intensity at 480 nm was large and decreased as the pH was increased (Figure 4B, □, left axis); this intensity showed a sigmoidal behavior, but we could not determine the pK_a of the titration due to the absence of an acidic baseline. Conversely, the $\langle\lambda\rangle$ showed a titration with a pK_a of 5.12 ± 0.09 (Figure 4B, ■, right axis), which was similar, within error, to that obtained by intrinsic fluorescence.

Far-UV CD. We used far-UV CD in the analysis of the unfolding of MJ0729 as a spectroscopic probe that is sensitive to protein secondary structure (*39*). The CD spectrum of MJ0729 at physiological pH and 25 °C exhibited a shape characteristic of helical proteins (Figure 5B). The estimated percentage of helical structure at 222 nm was 15%, lower than that determined by FTIR (see below). The differences are probably due to the presence of aromatic signals absorbing at this wavelength (*39*). The shape of the far-UV CD spectrum did change only at the extremes of the pHs explored (Figure 5A). The sigmoidal behavior of the ellipticity at 222 nm at acidic pH yielded a pK_a of 3.65 ± 0.06, different from those measured by fluorescence.

Near UV. We used near-UV CD to detect possible changes in the asymmetric environment of aromatic residues (*39*). The near-UV of MJ0729 at pH 7.0 was intense, with bands at 280 nm, which indicate the presence of an asymmetric environment for all the aromatic residues (two Phe residues, one Trp residue, and five Tyr residues per monomer) (Figure

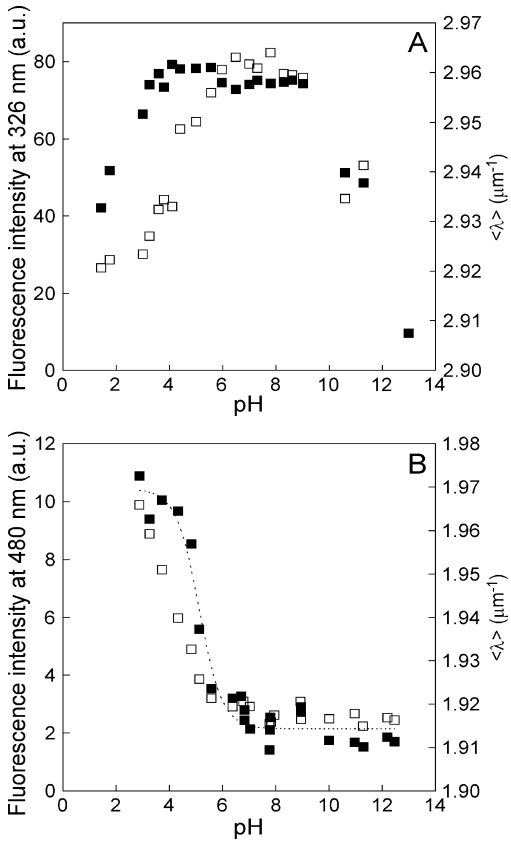

FIGURE 4: Fluorescence of MJ0729. (A) Intrinsic fluorescence. The average energy (right axis, ■) and the intensity at 326 nm (left axis, □) are represented vs the pH by excitation at 280 nm (similar results were observed by excitation at 295 nm). (B) ANS fluorescence. The average energy (right axis, ■) and the intensity at 480 nm (left axis, □) are represented vs the pH. The dotted line through the $\langle\lambda\rangle$ data is the fitting to eq 8. Experiments were conducted at 25 °C.

5C). At pH 2.0, the features of the near-UV spectrum disappeared (data not shown), indicating that the protein was unfolded.

FTIR Experiments. FTIR is a powerful method for the investigation of secondary and tertiary structures, by following spectroscopic changes of amide I′ and Tyr (appearing at 1515 cm^{-1}) bands, respectively. The main advantage in comparison with CD and fluorescence is that FTIR is much more sensitive to the presence of β structure or random coil. Structural information is obtained by the analysis of the amide I′ region of the spectrum (1700−1600 cm^{-1}). The absorbance of this band is mainly due to the stretching vibration of the carbonyl peptide bond, whose frequency is highly sensitive to hydrogen bonding and, thus, to protein secondary structure (*40*). The band of the FTIR spectra in MJ0729 at pH 7.0 was centered near 1644 cm^{-1}, which is characteristic of the presence of nonordered conformations. Fourier self-deconvolution and derivation analysis (see Experimental Procedures) exhibited maxima centred at 1687.7, 1677.2, 1667.1, 1656.4, 1646.5, 1636.4, 1628.3, and 1613.4 cm^{-1} (Table 2). The 1613.4 cm^{-1} component corresponds to tyrosine side chain vibrations (*41*), and the other maxima are assigned to vibration of groups involved in different secondary structural motifs. The 1628.3 and 1636.4 cm^{-1} bands are assigned to (π,0) β sheet structure intermolecular and β-sheet structure intramolecular, respectively. The 1646.5 cm^{-1} band is assigned to random coil structure. The 1656.4 cm^{-1} band is assigned to α helix or disordered structure. The 1667.1, 1677.2, and 1687.7 cm^{-1} components are assigned to turns and loops. These results are similar to the percentages of secondary structure of the protomer in the resolved tetrameric X-ray structure (21% of α-helical structure and 7% of β-sheet) (L. A. Martínez-Cruz, unpublished results).

At pH 2.5, the percentage of helical structure disappeared (Table 2) and most of the residual structure was involved in intermolecular β sheet, suggesting aggregation (the 1628 cm^{-1} band). These results support our previous fluorescence and CD findings, where the protein was unfolded at low pHs (Figures 4A and 5A). Furthermore, they suggest that populated species at highly acidic pHs (≤2.5) are high molecular-mass species formed by intermolecular hydrogen bonds. We could not carry out similar experiments at basic pHs, because the protein precipitated.

The amide II band (centred at 1547 cm^{-1}) results primarily from NH bending vibrations in the peptide backbone (*42*). Its residual intensity remaining after D_2O exchange arises from nonexchangeable NH groups, and therefore, it reports on the inaccessibility of the protein core to the solvent due to tertiary structure. We did not observe an amide II band at acidic pHs (data not shown). These facts suggest that all the NH protons were solvent-accessible, and then the structure was very flexible or disordered (as suggested by the other spectroscopic techniques).

Nuclear Magnetic Resonance. NMR can give information about the general fold of a polypeptide chain in solution at residue level. The upfield-shifted regions of the one-dimensional NMR spectra of MJ0729 at 25 °C and pH 5.0, 6.80, and 8.8 indicate that the protein was folded, with methyl groups appearing at chemical shifts as low as 0.2 ppm (Figure 1 of the Supporting Information). The signals were broader than expected from protein molecular size (14302.7 Da), suggesting the presence of conformational exchange, probably due to the oligomerization equilibria described. The fact that the spectra at different pHs are similar suggests that the protomer does not change its structure significantly between pH 5.0 and 9.0.

In summary, we conclude from the spectroscopic probes above that not only the quaternary structure but also the secondary and tertiary ones of MJ0729 change at the extremes of pH. Furthermore, the secondary and tertiary structures (as monitored by the tryptophan residues and NMR spectra) remained unaltered between pH 5.0 and 9.0; at pH 2.5, the secondary structure changed dramatically (Table 2), as shown by FTIR, due to the presence of aggregated species.

Thermal and Chemical Stability of MJ0729. We have described the conformational preferences of MJ0729 over a wide pH range, but we do not know how stable the protein is between pH 5.0 and 9.0, where the secondary and tertiary structures remained basically unaltered. Then, next, we tried to determine the stability of the protein in that pH range.

Thermal Denaturations Followed by Far and Near UV. The ellipticities at 222 nm (far UV) and 280 nm (near UV) did not change in a sigmoidal fashion (Figure 5D). The denaturation process was fully reversible at all explored pHs between 5.0 and 9.0.

Thermal Denaturations Followed by FTIR. The thermal dependence of the amide I′ band was used to assess the stability of the protein secondary structure. Thermal dena-

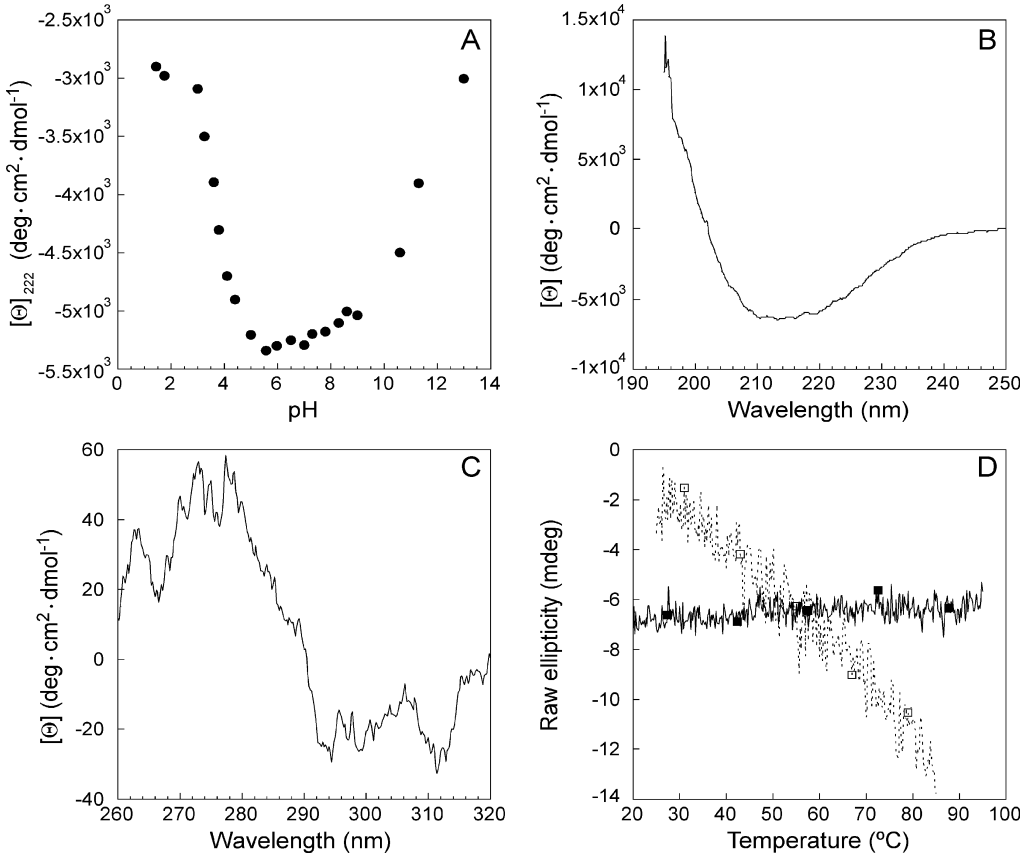

FIGURE 5: Far- and near-UV spectra of MJ0729. (A) Change in $[\theta]$ as the pH was modified. (B) Far-UV CD spectrum at pH 7.0 and 25 °C. (C) Near-UV spectrum at pH 7.0 and 25 ° C. (D) Thermal denaturations followed by the changes in ellipticity at 280 nm (□ and dotted lines) and 222 nm (■ and solid lines) at pH 7.0; the scale on the *y*-axis is arbitrary. The conditions were as follows: 10−15 μM protomer protein (far-UV) or 30−40 μM protomer protein (near-UV) with a buffer concentration of 10 mM in all cases.

Table 2: Secondary Structural Analysis of MJ0729 As Determined by FTIR at Two pHs and 25 °C[a]

wavenumber (cm^{-1})	structural assignment[b]	structure content (%)	
		pH 7.0	pH 2.5
1687.7	turns and loops	4	
1677.2	turns and loops	10	2
1667.1	turns and loops	6	
1656.4	α helix	26	6
1646.5	random coil	21	20
1636.4	β sheet intramolecular	17	
1628.3	β sheet intermolecular	14	72[b]
1613.4	tyrosine	2	

[a] Errors in the wavenumber are estimated to be ±2 cm^{-1}. [b] There were other bands at 1618.4 cm^{-1} at this pH whose assignment is due to the formation of β sheet intermolecular, and the sum of both components is indicated.

turation results in amide I′ band widening and other spectral changes, including the appearance of components at 1620 and 1684 cm^{-1}, which are related to aggregation of thermally denatured proteins (*40*).

The thermal denaturation of MJ0729 at pH 7.0 showed two transitions when the change in the tyrosine bands was followed (Figure 6A) and only one when the denaturation of the amide I′ band was observed (Figure 6B). Thermal unfolding of protein involves not only loss of secondary structure but also changes in the microenvironment of side chain groups. Aromatic ring stretching vibrations of tyrosine at 1515 cm^{-1} provide a specific local monitor for both tertiary and secondary structural changes (*41*). The T_m of the amide I′ band was the same as that obtained in the second thermal transition of the tyrosine bands (Figure 6A); this suggests that the second transition involved a global unfolding of MJ0729. The first transition probably implicated a local denaturation around the tryrosine amino acids. The thermal transitions were irreversible at pH 7.0, due to aggregation processes occurring at the high concentrations used in the FTIR experiments. Thermally induced changes in spectroscopic parameters were not observed at pH 2.5, then supporting the above observations that at this low pH the protein was unfolded (Figures 4A and 5A).

Thermal Denaturation Followed by DSC. The heat-induced denaturation of MJ0729 was also monitored by the change in the excess heat capacity of the protein solution when it was heated at a constant scan rate. At pH 7.0, the protein was highly stable upon heating, although the reversibility of the process was shown to be only marginal (only ∼30% of the endotherm was detected upon rescanning the protein at the lowest concentration tested, 0.1 mg/mL). The level of irreversibility shown in these experiments, which was not observed in the CD measurements (see above), is probably due to the much higher temperatures reached in the DSC experiments (125 °C).

Several experiments were performed under the same experimental conditions (see Experimental Procedures) at protein concentrations within the interval of 0.1−1 mg/mL (protein precipitation precluded the use of higher concentrations). The excess heat capacity functions obtained at different protein concentrations were almost identical (see the Supporting Information); they were characterized by a

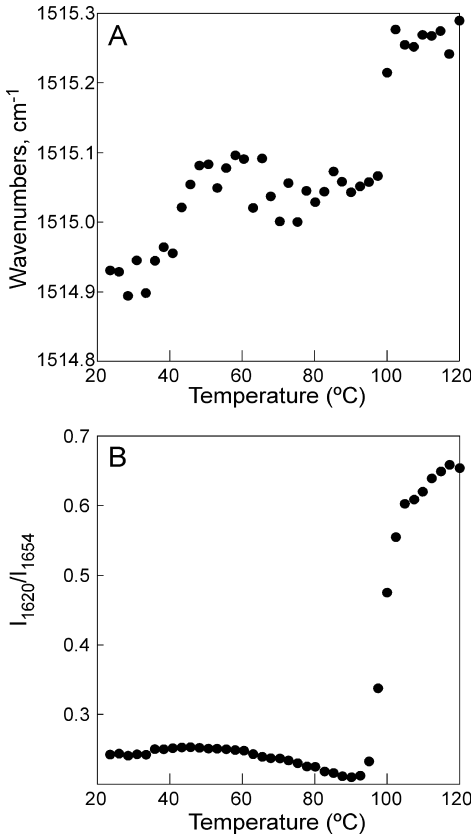

FIGURE 6: FTIR of MJ0729. Thermal denaturation of MJ0729 followed by the changes in the tyrosine band (1515 cm^{-1}) (A) and the rate of the intensities, I_{1620}/I_{1654} (B).

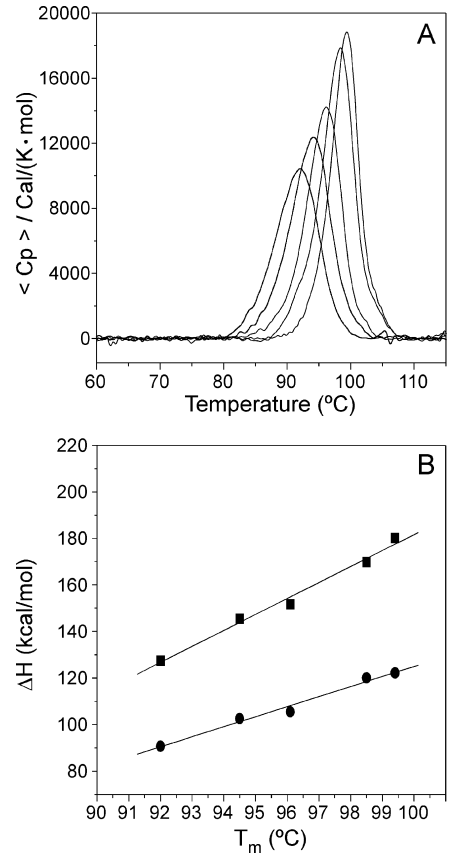

FIGURE 7: Thermal denaturation of MJ0729 in the presence of GdmCl at pH 7.0 monitored by DSC. (A) The protein, at a concentration of 0.2 mg/mL (14.0 μM monomer), was heated at a constant scanning rate (90 °C/h) to 125 °C in the presence of increasing concentrations of GdmCl (from right to left): 0, 0.5, 1.0, 1.5, and 2.0 M. (B) Temperature dependence of the van't Hoff (■) and calorimetric (●) enthalpy changes upon protein unfolding at different concentrations of denaturant. The slopes of the linear regressions yield the van't Hoff and calorimetric heat capacity changes upon unfolding, 6.8 and 4.3 kcal K^{-1} mol^{-1}, respectively.

single endotherm with its maximum located at the same temperature, 99.5 ± 0.1 °C (average of seven scans), and their calorimetric enthalpy showing no concentration dependency ($\Delta H = 122 \pm 12$ kcal/mol of monomer). At the lowest protein concentrations tested (from 0.1 to 0.3 mg/mL; see Figure 7A), the endotherm is symmetrical (with no evidence of dissociation coupled to the denaturation of the protein) and fits fairly well to the two-state unfolding mechanism. Under these conditions, the van't Hoff enthalpy change upon unfolding (180 kcal/mol) is substantially larger than the calorimetric one (122 kcal/mol), suggesting that the cooperative folding unit is larger than one monomer. On the other hand, the van't Hoff to calorimetric enthalpy ratio is lower than 2 (~1.5) which points to the incomplete unfolding of the protein dimer to yield a partial (still dimeric) state of the protein particularly prone to aggregation. As the concentration of the protein is increased, the contribution of the irreversible aggregation of the protein to the overall denaturation process becomes more evident since the endotherm becomes gradually more skewed toward its high-temperature side, together with a concomitant increase in the noise of the post-transitional baseline, and its deviation of its normal trend due to the contribution of an exothermic process (which is characteristic of protein precipitation).

Figure 7B also shows the dependence of both van't Hoff and calorimetric enthalpy changes on melting temperature, T_m, observed for each concentration of GdmCl within the range of 0−2 M. The presence of increasing concentrations of denaturant destabilizes the native state of the protein $[(\partial T_m)/(\partial[\text{GdmCl}]) = 3.78$ °C L/mol (see the Supporting Information)] without affecting the two-state unfolding mechanism described in aqueous buffer. The van't Hoff to calorimetric enthalpy ratio ranged from 1.35 at 2 M GdmCl to 1.5 in buffer. The heat capacity changes upon protein unfolding were determined from the slopes of the linear correlations in the ΔH versus T_m plots. As in the case of the enthalpy changes discussed above, the calorimetric heat capacity change upon unfolding amounts to 4.3 kcal K^{-1} mol^{-1}, significantly lower than the one calculated from the van't Hoff enthalpies, 6.8 kcal K^{-1} mol^{-1}. Therefore, we must conclude from the calorimetric data that the protein seems to be a dimer in its native state (below the T_m) and unfolds cooperatively to give a dimeric partially folded state that may become involved in the irreversible precipitation of the protein.

Chemical Denaturation. From the thermal experiments, it is clear that MJ0729 is a highly thermostable protein. Then, we carried out equilibrium chemical denaturations monitored by fluorescence at different temperatures at physiological pH where the protein was mainly a dimer (Figure 2A). The intensity of the fluorescence spectra at 326 nm and $\langle\lambda\rangle$ changed in a single sigmoidal fashion, as the concentration of GdmCl increased (Figure 8A). The fact that a single sigmoidal transition was observed suggests that dimer dissociation and monomer unfolding occurred concomitantly,

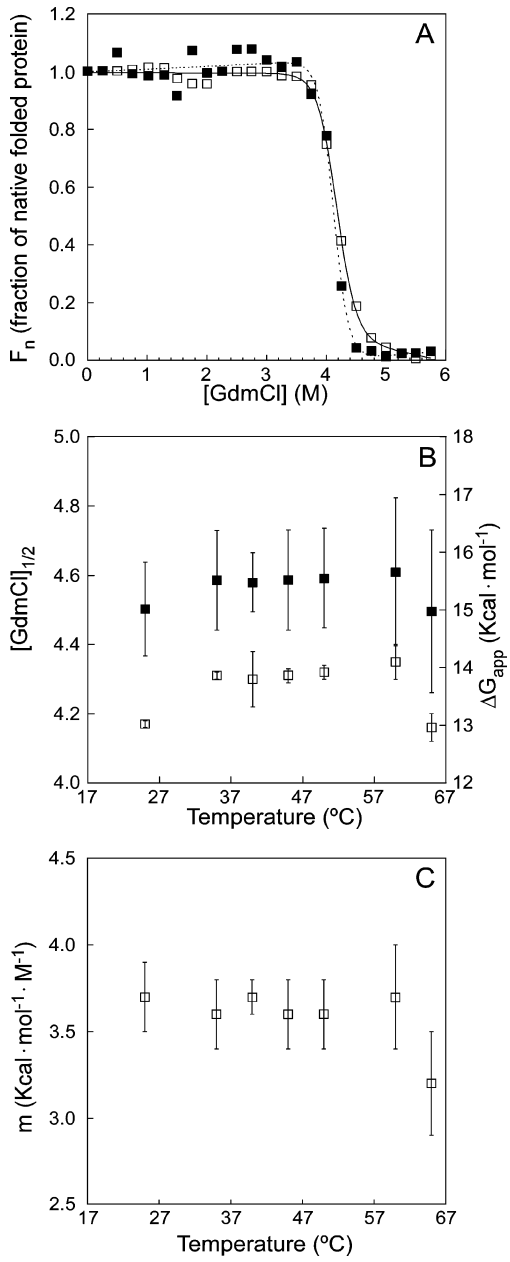

FIGURE 8: Chemical denaturation data of MJ0729. (A) Sigmoidal chemical denaturation data of 1.6 μM protomer of MJ0729 at 25 °C (□, solid line) and 65 °C (■, dotted line). The line is the fitting to a two-state model with protein concentration dependence. (B) Chemical denaturation midpoints [$[GdmCl]_{1/2}$ (□)] and apparent free energies [calculated as described by Park and Marqusee (*34*) (■)]. The error bars are errors from fitting to the two state-model (the ($[GdmCl]_{1/2}$) or from error propagation (ΔG_{app}). (C) *m* values of the chemical denaturations. Error bars are from fitting to the two-state model. The protein concentration was 1.26 μM protomer in all cases.

and then this transition should be concentration-dependent. To test that hypothesis, we carried out experiments at 25 °C, at two different concentrations: 1.26 and 1.54 μM protein (in protomer units). The $[GdmCl]_{1/2}$ values were 4.25 ± 0.02 and 4.30 ± 0.03 M, respectively, which are similar, but different enough to show the variation. Experiments at 4 μM were also conducted (Figure 4 of the Supporting Information), but the larger stability precluded an exact determination of the denaturation midpoint. The GdmCl denatutation experiments at any of the explored temperatures were reversible, and the $[GdmCl]_{1/2}$ decreased slightly within the temperature interval that was explored (Figure 8B). We should expect also a decrease in the *m* values at high temperatures as the protein becomes more unfolded; however, all the *m* values were similar, within error, in the temperature range that was explored (Figure 8C).

These results could suggest that the monomers are not sufficiently stable to exist isolated in solution. However, since they might be formed at high pH [this work (see above)], the monomers could be sufficiently stable to exist isolated in solution, but they are not significantly populated because dimer formation is highly favorable. Then, when solution conditions make the dimer dissociate, the monomeric species are no longer sufficiently stable to be populated.

DISCUSSION

The Oligomerization State of MJ0729 Is pH-Dependent. The main result of this work is that MJ0729 exhibits a pH-dependent oligomerization state: at pH < 2.5, the dominant species is a high-molecular mass species with intermolecular hydrogen bonding, as suggested by FTIR (Table 2 and Figure 2 of the Supporting Information). At higher pHs (between pH 4.5 and 5.3), these species dissociate to yield a tetrameric one [as shown by X-ray (Figure 1C) and suggested by the decreasing slope of the DLS curve (Figure 2A)]. It is important to bear in mind, however, that the measurements from DLS, as those of DOSY-NMR, are the result of the averaged properties of those species present in solution. Thus, the long R_s measured at low pH is the result of a *D* measured by DLS, which averages over the elongated native tetrameric species present, possibly other native-like tetrameric species, and the intermolecular hydrogen bonding species present at lower pHs [clearly shown by FTIR (Figure 2 of the Supporting Information)]. Thus, the R_s values at those low pHs are the result of different oligomeric species.

Oligomeric dissociation (tetrameric and the other oligomeric species) occurs because of titration of some aspartic and/or glutamic residues; further, this species probably dissociates to yield a highly oblong tetrameric species (see Results and Figure 2B). We have indirect evidence from the tetrameric state in solution, since the only confirmation comes from the presence of long-R_s species in DLS experiments (Figure 2A) and the calculations with SEDFIT (see Figure 2B). We can conclude that such a tetramer is not an artifact of the crystallization procedures or due to the high protein concentrations used, since it seems to be present in solution in a narrow pH range. At physiological pH, a dimeric species is present (Figure 2B), which, as suggested by DLS and DOSY measurements, also has an elongated shape. At pH >10, the protein either suffers rearrangements in the quaternary structure (which did not alter the dimeric state) or, alternatively, might dissociate to a monomeric species. It is important to indicate, however, that we do not have other alternative evidence of the presence of monomeric species at basic pHs, other than the DLS and DOSY measurements.

The secondary (CD) and tertiary (fluorescence) structures and the burial of hydrophobic residues (ANS) also were pH-dependent, but the apparent pK_a values were different; therefore, various residues were involved in the acquisition of the different types of structure. Interestingly enough, the acquisition of secondary structure occurs at very acidic pHs, with an apparent titration midpoint of 3.6 (Figure 5A),

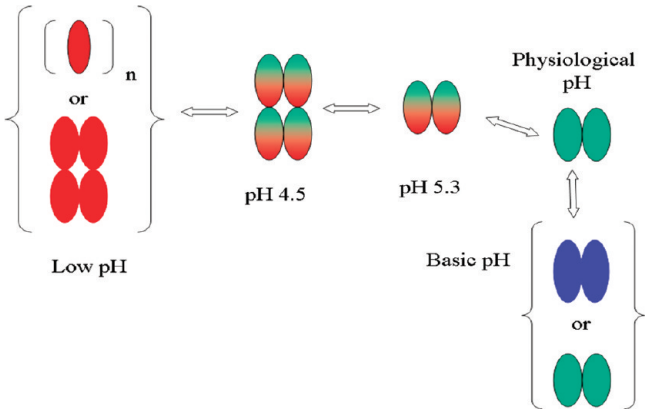

FIGURE 9: Model of association of MJ0729. The green oval indicates the monomer which has the whole hydrophobic surface buried. The red one indicates the monomer which has large patches of solvent-exposed hydrophobic surface. The red−green oval indicates species with some hydrophobic patches still solvent-exposed. At basic pH, a rearrangement of the dimeric form is suggested (in blue ovals), which could alter the native like structure of the dimer. Since the evidence of a monomeric species at this basic pH is weak, we have preferred, for the reader's sake, not to include it in this simplified pictorial model. The ovals indicate the elongated shape of the macromolecule.

suggesting that the residue(s) involved (probably aspartic and/or glutamic) is buried. On the other hand, the acquisition of native-like tertiary structure (Figure 4A) and the burial of hydrophobic residues (Figure 4B) occurred at pH ~5.0, similar to the pK_a of a solvent-exposed aspartic residue (*43*). These data suggest that the long-R_s species at low pH have native-like secondary structure, even though the tertiary structure is not fully attained. Then, MJ0729 seems to populate a molten globule state at low pH.

The formation of the quaternary native-like dimeric structure seems to be governed by the presence of residues with a pK_a close to 6.0 (Figure 2A); the only residue with such pK_a is a histidine (*43*), which is present in MJ0729 in the long loop connecting both CBS domains. However, this residue is far from the dimeric interface (L. A. Martínez-Cruz, unpublished results); then, the titration midpoint must be ascribed to an Asp and/or Glu residue, which should be hydrogen-bonded in the tetrameric species (hence, its high pK_a value).

Then, we can conclude that the fact that the pH titrations followed by different probes have different midpoints suggests that the pH denaturation process is not an "all-or-none" process. Thus, there are species with different secondary and tertiary structures, and even with a different amount of solvent-exposed hydrophobic surface.

However, the question of which molecular mechanisms are behind those changes in the quaternary structure can be raised. Since the changes are pH-dependent, they must rely on electrostatics. Furthermore, burial of hydrophobic surface must also be important; it has been recently argued that during formation of homooligomeric species, the largest interacting interface is maintained (*44*). In MJ0729, the total solvent accessible area buried upon dimer formation (from monomer) is 7653 Å^2, and the area buried per each dimer upon tetramer formation is 6133 Å^2 (*45*). Then, the most stable species seems to be the elongated dimeric one, and the tetramer can be envisioned as a dimer of dimers which keeps the elongated shape (Figure 9). The large surface area buried upon dimer formation also suggests that a well-folded MJ0729 monomer does not exist independently in solution.

There are protein families with members with different quaternary structure (*44, 46*). For instance, the members of the DsrEFH family can be monomeric, homodimeric, homotrimeric, and homohexameric; interestingly enough, as it happens with MJ0729, the proteins are not well-characterized, and some members of the family are thought to be involved in sulfur metabolism (*47, 48*). Changes in protein quaternary structure are associated with the evolution of new functions and regulation. A major mechanism in the evolution of protein function is the formation or disruption of interactions among the subunits forming a homooligomer. Recently, it has been shown that a switch from a monomeric helical protein to a dimeric one is based on a conversion of an α-helix to β-sheet, which forms the new dimer interface (*49*). At low pHs, in MJ0729, high-molecular mass oligomeric proteins are present, as shown by the presence of intermolecular β-sheets (Table 2 and Figure 2 of the Supporting Information). The titration of an acidic residue probably disrupts the intermolecular hydrogen bonds and changes the conformation of a β-strand to an α-helix conformation, making that oligomerization interfaces be formed by the docking of helices (Figure 1C). However, these helices in the tetrameric protein must be flickering and some of them highly unstable at these low pHs (as shown by the hydrogen exchange experiments followed by FTIR) due to solvent exposure of hydrophobic regions not yet fully buried [as indicated by the ANS experiments (Figure 4B)]. Then, titration of other acidic group(s) triggers a change to a dimeric form, which is the dominant species up to pH 9, as shown by ultracentrifugation experiments (Figure 2B).

High Equilibrium Stability of MJ0729. The relatively high thermal denaturation midpoints for MJ0729 are most likely related to its origin from a hyperthermophile organism. From a thermodynamic point of view, the thermal stability of a protein can be attained either by a large maximum in the free energy curve, the so-called ΔG_{max}, or by a low ΔC_p (or by both reasons together); either factor makes the free energy curve intersect the *x*-axis (the *T*-axis) at very high values. Calculation of the ΔG curve in MJ0729 from the equilibrium denaturation experiments (Figure 8B) indicates that the curve is very broad (probably due to a small ΔC_p), and then, the T_m is very high. In addition to these factors, MJ0729 self-associates, and according to the action mass, a higher protein concentration implies a greater conformational stabilization. For instance, high stability has been reported for the dimeric dihydrofolate reductase from *Thermotoga maritima* (*50*) and in the dimeric ORF56 from *Sulfolobus islandicus* (*51*). In the case of dihydrofolate reductase, the large stability has been attributed to an extremely slow unfolding rate. It is reasonable to assume that disruption of the large hydrophobic surface of dimeric MJ0729 must also be very slow.

An important question that has not yet been answered is what molecular mechanism(s) may underlie the very high stability of MJ0729. It has been suggested that a high level of quaternary structure is responsible for the extreme thermal stabilization of homooligomeric proteins (*52*). However, this hypothesis does not hold for MJ0729 since at low pH, where the tetrameric (and other oligomeric) species are present, the proteins exhibited irreversible thermal and chemical denaturations with smaller midpoints due to partial unfolding of

the protein (data not shown). High protein stability has also been attributed to a large number of salt bridges (*53*), and there are several salt bridges in the dimeric interface of MJ0729, which probably regulate the quaternary structure (see above).

On the other hand, a comparison among the sequences of proteins from thermophiles and mesophiles can shed some light on that raised question (*54*). Protein sequences from thermophiles show a large proportion of charged (Arg, Asp, Glu, and Lys) and hydrophobic (Gly, Ile, Pro, and Val) residues and, concomitantly, a low percentage of uncharged polar amino acids (Asn, Gln, and Thr). In MJ0729, the percentage of charged and hydrophobic accounts for 50% of the total number of residues, and then it is larger than that of uncharged ones (10%), suggesting that its sequence also determines its thermostability. More detailed molecular interpretations of the stability of oligomeric proteins are difficult. Although there is a great deal of research these days into prediction of subunit affinities (as shown by CAPRI results), accurate predictions remain elusive (*55*). We can only hypothesize that the high stability of the oligomeric MJ0729 is important for its function, as described recently for other oligomeric proteins (*56*), although the exact function of MJ0729 remains to be elucidated.

Equilibrium Unfolding of MJ0729. Most of the previous work on protein folding has focused on small single-domain proteins that fold rapidly and avoid aggregation (*57*). However, there are many proteins in the cell which are large, with multiple folding domains and/or subunits and which probably do not follow the simple folding principles established with results from small proteins. Further complications arise in multimeric proteins, where their folding involves not only intramolecular interactions but also intermolecular ones. How the amino acid sequence controls subunit−subunit interactions and which additional stability is conferred by such contacts can be only addressed by characterizing the stability and conformational properties of multimeric proteins.

The stability of other proteins (that is, free energy of unfolding plus energy of dissociation) is in the range of 15−35 kcal/mol (*35, 58, 59*); for those proteins [most of them dimeric (*59*)] showing a single transition, the conformational stabilities range from 10 to 27 kcal/mol. These values are within the range of the apparent free energy (*34*) determined in MJ0729 (Figure 8B). Thus, MJ0729 behaves, from the point of view of stability, as the majority of the oligomeric proteins reported to date.

Biological Implications of the Different Oligomerization States of MJ0729. Protein self-association is a commom process, and in fact, approximately 50−70% of the known structures assemble into homomers (*44*). Oligomerization improves protein stability, controls the accessibility and specificity of particular sites, and restricts the presence of enzyme activity. Even in virus-encoded proteinases self-association can restrict, in time, infection stages. On the other hand, mechanisms preventing oligomerization through intramolecular interactions have been described in some proteins. For instance, members of the Wiskott-Aldrich syndrome protein family form intramolecular interactions to hide otherwise protein interacting regions (*60*); furthermore, the cysteine proteinase of hepatitis A virus forms homodimers in the millimolar range, which is reduced to micromolar when the viral RNA is present (*61*). Thus, these examples indicate that self-association is important and can contribute to regulation of some physiological phenomena.

In this work, we have shown that the two-CBS domain protein MJ0729 forms several pH-dependent homomeric species. It could be thought that oligomerization is a result of the particular domain dissection we have chosen to study. There are, however, several pieces of evidence that suggest that the self-association is intrinsic and specific to the whole polypeptide chain: (i) oligomerization could be measured by four different techniques (Figures 2 and 3 and Table 1), and (ii) the self-oligomerization seems to be specific by burying particular patches of hydrophobic surfaces (Figure 4B). Oligomerization, in addition, improves the stability of MJ0729, as shown by chemical denaturation experiments carried out at different protomer concentrations (Figure 4 of the Supporting Information). Further, CBS domains have been described as being involved in the multimerization of the homocystathionine β-synthase (*3*). Protein engineering studies are necessary to find which residues are responsible for modulating the oligomerization states of MJ0729 and whether this modulation can be conducted by the binding of natural substrates, such as adenosyl groups, or even by binding to other proteins.

ACKNOWLEDGMENT

We deeply thank both reviewers for helpful comments, insights, and suggestions. We deeply thank May García, María del Carmen Fuster, María T. Garzón, and Javier Casanova for excellent technical assistance.

SUPPORTING INFORMATION AVAILABLE

NMR spectra of MJ0729 at different pHs (Figure 1), FTIR spectra at different temperatures at two pHs (Figure 2), excess heat capacity change of MJ0729 versus temperature for different protein concentrations within the range of 0.1−0.9 mg/mL (Figure 3A) and showing the dependence of T_m on the concentration of GdmCl (Figure 3B), and sigmoidal chemical denaturation data of MJ0729 protomer at 1.6 (□) and 4 μM (■) at 25 °C (Figure 4). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bateman, A. (1997) The structure of a domain common to archaebacteria and the homocystinuria disease protein. *Trends Biochem. Sci. 22*, 12–13.
2. Zhang, R., Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., Joachimiak, A., and Collart, F. R. (1999) Characteristics and crystal structure of bacterial inosine-5′-monophosphate dehydrogenase. *Biochemistry 38*, 4691–4700.
3. Ignoul, S., and Eggermont, J. (2005) CBS domains: Structure, function and pathology in human proteins. *Am. J. Physiol. 289*, C1369–C1378.
4. Estévez, R., Schroeder, B. C., Accardi, A., Jentsch, T. J., and Pusch, M. (2003) Conservation of chloride channel structure revealed by an inhibitor binding site in ClC-1. *Neuron 38*, 47–59.
5. Kemp, J. (2004) Bateman domains and adenosine derivatives form a binding contract. *J. Clin. Invest. 113*, 182–184.
6. Rudolph, M. J., Amodeo, G. A., Iram, S. H., Hong, S. P., Pirino, G., Carlson, M., and Tong, L. (2007) Structure of the Bateman2 domain of yeast Snf4: Dimeric association and relevance for AMP binding. *Structure 15*, 65–74.
7. Meyer, S., Savaresi, S., Forster, I. C., and Dutzler, R. (2007) Nucleotide recognition by the cytoplasmic domain of the human chloride transporter ClC-5. *Nat. Struct. Mol. Biol. 14*, 60–67. [Erratum: (2007) *Nat. Struct. Mol. Biol. 14*, 172].

8. Amodeo, G. A., Rudolph, M. J., and Tong, L. (2007) Crystal structure of the heterotrimer core of *Saccharomyces cerevisiae* AMPK homologue SNF. *Nature 449*, 492–495.
9. Day, P., Sharff, A., Parra, L., Cleasby, A., Williams, M., Hörer, S., Nar, H., Redemann, N., Tickle, I., and Yon, J. (2007) Structure of a CBS-domain pair from the regulatory gamma1 subunit of human AMPK in complex with AMP and ZMP. *Acta Crystallogr. D63*, 587–596.
10. Jin, X., Townley, R., and Shapiro, L. (2007) Structural insight into AMPK regulation: ADP comes into play. *Structure 15*, 1285–1295.
11. Townley, R., and Shapiro, L. (2007) Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science 315*, 1726–1729.
12. Xiao, B., Heath, R., Saiu, P., Leiper, F. C., Leone, P., Jing, C., Walker, P. A., Haire, L., Eccleston, J. F., Davis, C. T., Martin, S. R., Carling, D., and Gamblin, S. J. (2007) Structural basis for AMP binding to mammalian AMP-activated protein kinase. *Nature 449*, 496–500.
13. Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D., Richardson, D. L., Kerlavage, A. R., Graham, D. E., Kyrpides, N. C., Fleischmann, R. D., Quackenbush, J., Lee, N. H., Sutton, G. G., Gill, S., Kirkness, E. F., Dougherty, B. A., McKenney, K., Adams, M. D., Loftus, B., Peterson, S., Reich, C. I., McNeil, L. K., Badger, J. H., Glodek, A., Zhou, L., Overbeek, R., Gocayne, J. D., Weidman, J. F., McDonald, L., Utterback, T., Cotton, M. D., Spriggs, T., Artiach, P., Kaine, B. P., Sykes, S. M., Sadow, P. W., D'Andrea, K. P., Bowman, C., Fujii, C., Garland, S. A., Mason, T. M., Olsen, G. J., Fraser, C. M., Smith, H. O., Woese, C. R., and Venter, J. C. (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature 390*, 364–370. [Erratum: (1998) *Nature 394*, 101].
14. Zhao, T., Cruz, F., and Ferry, J. G. (2001) Iron-sulfur flavoprotein (Isf) from *Methanosarcina thermophila* is the prototype of a widely distributed family. *J. Bacteriol. 183*, 6225–6233.
15. Fernández-Millán, P., Kortazar, D., Lucas, M., Martínez-Chantar, M. L., Astigarraga, E., Fernández, J. A., Sabas, O., Albert, A., Mato, J. M., and Martínez-Cruz, L. A. (2008) Crystallization and preliminary crystallographic analysis of merohedrally twinned crystals of MJ0729, a CBS-domain Protein from *Methanocaldococcus jannaschii*. *Acta Crystallogr. F64*, 605–609.
16. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol. 131*, 266–280.
17. Pace, C. N., and Scholtz, J. M. (1997) Measuring the conformational stability of a protein. In *Protein Structure* (Creighton, T. E., Ed.) 2nd ed., pp 253−259, Oxford University Press, Oxford, U.K.
18. Muro-Pastor, M. I., Barrera, F. N., Reyes, J. C., Florencio, F. J., and Neira, J. L. (2003) The inactivating factor of glutamine synthetase, IF7, is a "natively unfolded" protein. *Protein Sci. 12*, 1443–1454.
19. Echabe, I., Encinar, J. A., and Arrondo, J. L. R. (1997) Effects of noise on the obtention of protein structure by band decomposition of the infrared spectrum. *Biospectroscopy 3*, 469–475.
20. Encinar, J. A., Mallo, G. V., Mizyrycki, C., Giono, L., Gonzalez-Ros, J. M., Rico, M., Canepa, E., Moreno, S., Neira, J. L., and Iovanna, J. L. (2001) Human p8 is a HMG-I/Y-like protein with DNA binding activity enhanced by phosphorylation. *J. Biol. Chem. 276*, 2742–2751.
21. Piotto, M., Saudek, V., and Sklenar, V. (1993) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR 2*, 661–665.
22. Dosset, P., Hus, J. C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR 16*, 23–28.
23. Price, W. S. (1997) Pulse-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: Part I. Basic theory. *Concepts Magn. Reson. 9*, 299–336.
24. Lapham, J., Rife, J. P., Moore, P. B., and Crothers, D. M. (1997) Measurement of diffusion constants for nucleic acids by NMR. *J. Biomol. NMR 10*, 255–262.
25. Hinkle, A., Goranson, A., Butters, C. A., and Tobacman, L. S. (1999) Roles for the troponin tail domain in thin filament assembly and regulation. A deletional study of cardiac troponin T. *J. Biol. Chem. 274*, 7157–7164.
26. Creighton, T. E. (1993) *Proteins. Structures and macromolecular properties*, 2nd ed., W. H. Freeman, New York.
27. Barth, I. L. (1984) *Modern methods of particle size analysis*, Wiley-Interscience, New York.
28. Koppel, D. E. (1972) Analysis of macromolecular polydispersity in intensity correlation spectroscopy: The method of cummulants. *J. Chem. Phys. 37*, 4814–4820.
29. Berne, P. J., and Pecora, R. J. (1976) *Dynamic light scattering with applications to chemistry, biology and physics*, Wiley-Interscience, New York.
30. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling. *Biophys. J. 78*, 1606–1619.
31. Laue, T. M. S., Ridgeway, T. M., and Pelletier, S. L. (1992) *Computer-aided interpretation of analytical sedimentation data for proteins*, Royal Society of Chemistry, Cambridge, U.K.
32. Royer, C. A. (1995) Fluorescence spectroscopy. *Protein Stability and Folding* (Shirley, B. A., Ed.) pp 65−89, Humana Press, Totowa, NJ.
33. Backmann, J., Schäfer, G., Wyns, L., and Bönisch, H. (1998) Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon *Sulfolobus acidocaldarius*. *J. Mol. Biol. 284*, 817–833.
34. Park, C., and Marqusee, S. (2004) Analysis of the stability of multimeric proteins by effective ΔG and effective *m*-values. *Protein Sci. 13*, 2553–2558.
35. Rumfeldt, J. A. O., Stathopulos, P. B., Chakaravarty, A., Lepock, J. R., and Meiering, E. M. (2006) Mechanism and thermodynamics of guanidinium chloride-induced denaturation of ALS-associated mutant Cu,Zn superoxide dismutases. *J. Mol. Biol. 355*, 106–123.
36. Schuck, P. (2006) Diffusion deconvoluted sedimentation coefficient distributions for the analysis of interacting and non-interacting protein mixtures. In *Analytical ultracentrifugation (techniques and methods)* (Scott, D. J., Harding, S. E., and Row, A. J., Eds.) pp 15−40, Royal Society of Chemistry, London.
37. Lebowitz, J., Lewis, M. C., and Schuck, P. (2006) Modern analytical ultracentrifugation in protein science: A tutorial review. *Protein Sci. 11*, 2067–2079.
38. Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., and Gilmanshin, R. I. (1991) Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers 31*, 119–128.
39. Kelly, S. M., and Price, N. C. (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci. 1*, 349–384.
40. Surewicz, W. K., and Mantsch, H. H. (1988) New insight into protein secondary structure from resolution-enhanced infrared spectra. *Biochim. Biophys. Acta 952*, 115–130.
41. Fabian, H., and and Mantsch, H. H. (1995) Ribonuclease A revisited: Infrared spectroscopic evidence for lack of native-like secondary structures in the thermally denatured state. *Biochemistry 34*, 13651–13655.
42. van Stokkum, I. H., Linsdell, H., Hadden, J. M., Haris, P. I., Chapman, D., and Bloemendal, M. (1995) Temperature-induced changes in protein structures studied by Fourier transform infrared spectroscopy and global analysis. *Biochemistry 34*, 10508–10518.
43. Thurlkill, R. L., Grimsley, G. R., Scholtz, J. M., and Pace, C. N. (2006) p*K* values of the ionizable groups of proteins. *Protein Sci. 15*, 1214–1218.
44. Levy, E. D., Erba, E. B., Robinson, C. V., and Teichmann, S. A. (2008) Assembly reflects evolution of protein complexes. *Nature 453*, 1262–1265.
45. Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R. F., Sykes, B. D., and Wishart, D. S. (2003) VADAR: A web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res. 31*, 3316–3319.
46. Ponstingl, H., Kabir, T., Gorse, D., and Thornton, J. M. (2005) Morphological aspects of oligomeric protein structures. *Prog. Biophys. Mol. Biol. 89*, 9–35.
47. Gaspar, J. A., Liu, C., Vassall, K. A., Meglei, G., Stephen, R., Stathopulos, P. B., Pineda-Lucena, A., Wu, B., Yee, A., Arrowsmith, C. H., and Meiering, E. M. (2005) A novel member of the YchN-like fold: Solution structure of the hypothetical protein Tm0979 from *Thermotoga maritima*. *Protein Sci. 14*, 216–223.
48. Dahl, C., Engels, S., Pott-Sperling, A. S., Schulte, A., Sander, J., Lubbe, Y., Deuster, O., and Brune, D. C. (2005) Novel genes of the dsr gene cluster and evidence for close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium *Allochromatium vinosum*. *J. Bacteriol. 187*, 1392–1404.
49. Roessler, C. G., Hall, B. M., Anderson, W. J., Ingram, W. M., Roberts, S. A., Montfort, W. R., and Cordes, M. H. (2008)

Transitive homology-guided structural studies lead to discovery of Cro proteins with 40% sequence identity but different folds. *Proc. Natl. Acad. Sci. U.S.A. 105*, 2343–2348.

50. Dams, T., and Jaenicke, R. (1999) Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochemistry 38*, 9169–9178.
51. Zeeb, M., Lipps, G., Lilie, H., and Balbach, J. (2004) Folding and association of an extremely stable dimeric protein from *Sulfolobus islandicus*. *J. Mol. Biol. 336*, 227–240.
52. Wang, W. C., Chiu, W. C., Hsu, S. K., Wu, C. L., Chen, C. Y., Liu, J. S., and Hsu, W. H. (2004) Structural basis for catalytic racemization and substrate specificity of an N-acylamino acid racemase homologue from *Deinococcus radiodurans*. *J. Mol. Biol. 342*, 155–169.
53. Razvi, A., and Scholtz, J. M. (2006) Lessons in stability from thermophilic proteins. *Protein Sci. 15*, 1569–1578.
54. Chakravarty, S., and Varadarajan, R. (2000) Elucidation of determinants of protein stability through genome sequence. *FEBS Lett. 470*, 65–69.
55. Lensink, M. F., Mendez, R., and Wodak, S. J. (2007) Docking and scoring protein complexes: CAPRI 3rd Edition. *Proteins 69*, 704–718.
56. Franco, J., Bañuelos, S., Falces, J., Muga, A., and Urbaneja, M. A. (2008) Thermodynamic characterization of nucleopasmin unfolding: Interplay between function and stability. *Biochemistry 47*, 7954–7962.
57. Jackson, S. E. (1998) How do small single-domain proteins fold? *Folding Des. 3*, R81–R91.
58. Sinha, S., Mitra, N., Kumar, G., Bajaj, K., and Surolia, A. (2005) Unfolding studies of soybean agglutinin and concanavalin A tetramers: A comparative account. *Biophys. J. 88*, 1300–1310.
59. Neet, K. E., and Timm, D. E. (1994) Conformational stability of dimeric proteins: Quantitative studies by equilibrium denaturation. *Protein Sci. 3*, 2167–2174.
60. Millard, T. H., Sharp, S. J., and Machensky, L. M. (2004) Signalling to actin assembly via the WASP (Wiskott-Aldrich syndrome protein) family proteins and the Arp2/3 complex. *Biochem. J. 380*, 1–17.
61. Peters, H., Kusov, Y. Y., Meyer, S., Bernie, A. J., Baumi, E., Wolff, M., Rademacher, C., Peters, T., and Grauss-Muller, V. (2005) Hepatitis A virus proteinase 3C binding to viral RNA: Correlation with substrate binding and enzyme dimerization. *Biochem. J. 385*, 363–370.

BI801920R