

In search for globally disordered apo-parvalbumins: Case of parvalbumin β -1 from coho salmon



Alisa A. Vologzhannikova^{a,*}, Polina A. Khorn^a, Alexei S. Kazakov^a, Ramis G. Ismailov^a,
Andrei S. Sokolov^a, Vladimir N. Uversky^{a,b,**}, Eugene A. Permyakov^{a,c}, Sergei E. Permyakov^{a,c,*}

^a Institute for Biological Instrumentation of the Russian Academy of Sciences, Pushchino, Moscow region, 142290, Russia

^b Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, 33647, USA

^c Pushchino State Institute of Natural Sciences, Pushchino, Moscow region, 142290, Russia

ARTICLE INFO

Keywords:

Protein stability
Intrinsically disordered proteins
Molten globule state
Calcium-binding proteins
EF-hand
Parvalbumin
Oncomodulin

ABSTRACT

Parvalbumin (PA) is a classical EF-hand calcium-binding protein of muscle, neuronal, and other tissues, and a major fish allergen. Although certain apo-PAs lack tertiary structure, functional implications of that feature and its structural prerequisites remain unclear. In a search for unstable PAs, we probed conformational stability of parvalbumin β -1 from coho salmon (csPA), a cold water fish species, using circular dichroism, scanning calorimetry, hydrophobic probe fluorescence, limited proteolysis, chemical crosslinking and dynamic light scattering techniques. Apo-csPA is shown to be mainly monomeric protein with markedly disorganized secondary structure and lack of rigid tertiary structure. Examination of per-residue propensity for intrinsic disorder in the PA groups with either folded or unfolded apo-form using the average PONDR^V VSL2P profiles revealed that the N-terminal region that includes α -helix A, AB-loop and N-terminal half of α -helix B is predicted to be less ordered in PAs with disordered apo-state. Application of the structural criteria developed for discrimination of disordered PAs indicate that the latter comprise about 16–19% of all PAs. We show that structural instability of apo- β -PA serves as a hallmark of elevated calcium affinity of the protein. Therefore, the successful predictions of unstable apo-PAs might facilitate search for PAs with maximal calcium affinity and possibly serving as calcium sensors.

1. Introduction

Parvalbumin (α isoform) or oncomodulin (β isoform) (will refer to as 'PA') [1–4] is a mainly cytosolic Ca^{2+} -binding protein of vertebrates (reviewed in [5–7]) with two active $\text{Ca}^{2+}/\text{Mg}^{2+}$ -binding motifs of the EF-hand type (CD and EF domains, each containing a 12-residue metal-coordinating loop flanked by α -helices) merged into a four-helix bundle domain, which is covered by an inactive AB domain [8]. The α and β isoforms differ in isoelectric point (α : pI > 5; β : pI < 5) and in preference for certain amino acids in specific positions. For example, the presence of Cys in AB-loop and the lack of extra C-terminal residue in C-

terminal helix are inherent to β -PAs. PAs are characterized by the highest among Ca^{2+} -binding proteins affinities for Ca^{2+} and Mg^{2+} (equilibrium association constants reach 10^{10} M^{-1} and 10^6 M^{-1} , respectively [5,9]) that enable involvement of these proteins into muscle relaxation [10–14]. PA forms with altered $\text{Ca}^{2+}/\text{Mg}^{2+}$ selectivity were suggested for correction of heart diseases [15–18]. The outstanding structural stability of Ca^{2+} -bound PA [5] favors its high allergenicity [19]: PA is a major and cross-reactive fish allergen [20,21]. Hypoallergenic PA forms with impaired metal affinity are under trials for immunotherapy of fish allergy [22]. PA is often expressed in tumors and considered as a reliable marker of chromophobe renal cell

Abbreviations: IDP, intrinsically disordered protein; PA, parvalbumin/oncomodulin; intact PA, parvalbumin isolated from muscles; rWT PA, recombinant wild-type parvalbumin; α -PA, α isoform of parvalbumin; β -PA, β isoform of parvalbumin; csPA, coho salmon PA; apo-csPA, metal-free form of coho salmon PA, prepared as described in *Materials and methods* chapter; IPTG, isopropyl β -D-1-thiogalactopyranoside; PMSF, phenylmethanesulfonyl fluoride; 2-ME, 2-mercaptoethanol; DTT, DL-dithiothreitol; DTNB, 5,5'-dithiobis(2-nitrobenzoic acid); EDTA, ethylenediaminetetraacetic acid; TCA, trichloroacetic acid; Tris, tris(hydroxymethyl)aminomethane; HEPES, N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid); MOPS, 3-(N-morpholino)propanesulfonic acid; bis-ANS, 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid; GdmCl, guanidine hydrochloride; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; CD, circular dichroism spectroscopy; DSC, differential scanning calorimetry; DLS, dynamic light scattering; SI, sequence identity; CH-plots, charge-hydrophobicity plots; CDF, Cumulative Distribution Function

* Corresponding authors at: Institute for Biological Instrumentation of the Russian Academy of Sciences, Institutskaya str., 7, Pushchino, Moscow region, 142290, Russia

** Corresponding author at: Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, 12901 Bruce B. Downs Blvd., MDC07, Tampa, FL, 33647, USA.

E-mail addresses: lisiks.av@gmail.com (A.A. Vologzhannikova), vuversky@health.usf.edu (V.N. Uversky), permyakov.s@gmail.com (S.E. Permyakov).

<http://dx.doi.org/10.1016/j.ceca.2017.08.011>

Received 11 May 2017; Received in revised form 8 August 2017; Accepted 31 August 2017

Available online 07 September 2017

0143-4160/ © 2017 Elsevier Ltd. All rights reserved.

carcinoma and benign oncocyoma [23]. PA expressed in distal convoluted tubule affects handling of electrolytes in the kidneys [23], while β -PA produced by cochlear outer hair cells is critical for cochlear function [24]. Hormone-like activity has been reported for chicken β -PA, which is expressed by epithelial cells in thymus and circulate in the blood, stimulating development of cell-mediated immunity (reviewed in [25]). Extracellular β -PA is a potent growth factor for mature optic nerve and peripheral sensory neurons [26]. Furthermore, PAs found in cutaneous mucus of several frog and fish genera participate in innate bacterial defense and serve as chemoattractants for thamnophiine snakes [27]. Although exact role of PA in neuronal tissue is still enigmatic, PA-positive inhibitory interneurons maintain excitatory/inhibitory balance and high-frequency neuronal synchronization, and they are impaired in schizophrenia, epilepsy, bipolar disorder, some forms of autism, Alzheimer's disease, Huntington's disease and multiple sclerosis [28–37]. Specific PA-positive neurons are engaged in the control of attention [38] and sleep [39], long-term memory consolidation [40] and triggering of fear responses [26,41].

While Ca^{2+} -bound PAs exhibit high resistance to thermal/chemical denaturation and enzymatic proteolysis, metal depletion dramatically destabilizes their structure [5]. For example, Ca^{2+} -loaded pike α -PA unfolds with mid-transition temperature of 120 °C [42], while its urea-induced unfolding occurs at urea concentrations above 8 M [43]. Meanwhile, apo-form of pike α -PA lacks rigid tertiary structure and resembles 'molten globule' state [44], which classifies it as an 'intrinsically disordered protein' (IDP) [45–48]. Unbound states of IDPs fully or partially lack unique tertiary structure under physiological conditions in vitro. They exist as dynamic ensembles of mutually transmuting conformations, in which the polypeptide chains undergo non-cooperative conformational changes over a broad range of time scales. IDPs exert their function via interaction with specific partners, such as small cofactors, nucleic acids, lipids or other proteins, which often converts IDPs to more structured states. The structural flexibility of IDPs and their target binding-induced folding enables involvement of IDPs into regulation of numerous biological activities both in prokaryotes and eukaryotes. The database of experimentally validated IDPs, DisProt (<http://www.disprot.org/> [49]), currently includes at least 800 proteins (v0.5 11-05-2017).

The instability of apo-PA and its exceptional metal-induced structural stabilization [5] can be rationalized by charge-charge interactions: apo-PA is destabilized due to the highly unfavorable electrostatic repulsion of the nearby carboxylates within Ca^{2+} -binding loops of PA, which turns into favorable electrostatic attraction upon the metal association [44]. The unfolding of apo-PA ensures high solvent entropy gain upon the metal binding-induced folding of the protein [50,51], thereby favoring high metal affinity (partly discussed in [52]) of PA, which is vital for its functioning. Therefore, structural stability of apo-form of PA likely directly affects its functional properties. Another possibility is an exposure of the unfolded polypeptide regions of apo-PA for interaction with some targets, as observed for other IDPs. This phenomenon could reconsider the view that PA is a pure calcium buffer (although some extracellular activities of β -PA point out its calcium sensor function) [25].

In fact, unfolding of apo-PA is firmly established only for two PAs, α and β isoforms of northern pike PA [44]. The experimental data reported for baltic cod β -PA [53] suggest that its apo-form is also mostly unfolded. In other cases, apo-PAs exhibit rigid tertiary structure with lowered compared to the Ca^{2+} -bound state stability. For example, apo-forms of common carp β -PA [54], α and β isoforms of rat PA [9,55], and chicken β -PAs [56,57] possess cooperative transitions with mid-transition temperatures from 30 °C [9] to 66 °C [55], depending upon species, buffer ionic strength and composition. Meanwhile, incidence of unfolded apo-forms among PAs and structural origins of this phenomenon remain unclear. Discovery of novel members of PA group with unfolded/intrinsically disordered apo-form could provide a basis for in-depth analysis of structural traits of PAs crucial for their stability.

Since the reported unstable PAs (see above) are found exclusively within cold-water fish species, whereas more stable PAs mostly originate from warm-blooded species, intrinsically disordered apo-PAs likely should be searched in cold-blooded species adapted to living in cold environment. In this sense, coho or silver salmon (*Oncorhynchus kisutch*) is a proper choice due to the fact that the species is characterized by one of the lowest median temperature preferences of 4 °C [58]. To validate this hypothesis, parvalbumin β -1 from coho salmon (csPA) is studied here with regard to the structural properties of its apo-state. The experimental data generated by several physical and biochemical techniques univocally evidence that apo-csPA is a globally disordered protein. The structural analysis of all PAs shown to be unfolded in apo-state reveals their structural signatures, which were applied for prediction of incidence of unfolded apo-PAs. The predicted unstable PAs provide a basis for search of PAs with maximal calcium affinity and likely possessing calcium sensor functions.

2. Materials and methods

2.1. Materials

Molecular biology grade HEPES and buffer grade MOPS were purchased from Calbiochem and AppliChem, respectively. Ultra-grade Tris and analytical grade PMSF were from Amresco. Biochemistry grade sodium acetate and biotechnology grade DTT were bought from Diam (Moscow, Russia). Ultra-grade H_3BO_3 and glycine, standard solution of calcium chloride, EDTA, analytical grade DTNB, bis-ANS, formic acid, horse heart myoglobin, premium grade bovine α -chymotrypsin and Phenyl Sepharose 6 Fast Flow were from Sigma-Aldrich Co. Biochemistry grade guanidinium chloride was a product of Merck. USP grade sodium chloride, biotechnology grade 2-mercaptoethanol and molecular mass markers for SDS-PAGE were from Helicon (Moscow, Russia). Fura-2 was bought from Molecular Probes. HPLC grade acetonitrile was from Panreac Química S.L.U. SP Sepharose Fast Flow, Sephadex G-25 and PD midiTrap G-25/Sephacryl S-100 HR were products of Amersham Biosciences Co., Pharmacia LKB and GE Healthcare, respectively. Other chemicals were reagent grade or higher. All buffers and other solutions were prepared using ultrapure water (Millipore Simplicity 185 system). Plastics or quartz ware were used instead of glassware, to avoid contamination of protein samples with Ca^{2+} . DTT solutions were freshly prepared from degassed buffers immediately prior to the usage to avoid DTT oxidation. Thermo SnakeSkin dialysis tubing (3.5 kDa MWCO) and Millipore Amicon Ultra centrifugal filters (3.0 kDa MWCO) were used for dialysis and concentration of protein solutions, respectively.

2.2. Methods

2.2.1. Preparation of standard solution of EDTA potassium salt

0.1 M stock solution of EDTA potassium salt was prepared by EDTA weighting and its dissolution in deionized water by adjusting pH to neutral values. Accurate concentration of the EDTA solution was determined using CaCl_2 standard solution. 2 mM EDTA solution in 50 mM MOPS-KOH, 100 mM NaCl, pH 7.2 buffer was prepared from the EDTA stock solution, followed by its titration by the CaCl_2 solutions prepared from 0.1 M CaCl_2 standard solution in the presence of 3 μM Fura-2, a calcium-sensitive fluorescent probe. The calcium concentrations exceeding EDTA concentration in the solution were accompanied by increase in free calcium concentration and the respective changes in fluorescence emission of Fura-2, detected using Cary Eclipse spectrofluorimeter, Varian Inc. (excitation at 340 nm).

2.2.2. Isolation of recombinant coho salmon parvalbumin

The amino acid sequence of parvalbumin β -1 from *Oncorhynchus kisutch* (common name coho or silver salmon; NCBI Taxonomy Browser ID: 8019), an anadromous fish of the salmon family, was extracted from

TrEMBL entry E1UJ20. The DNA sequence was codon optimized for expression in *E. coli* using EMBOSS Backtranambig service of EMBL-EBI [59]. The ORF of coho salmon PA (csPA) was cloned into pET-29a(+) vector (Novagen®) between the NdeI and XhoI restriction sites. Nucleotide sequence of the PA gene was verified by automatic DNA sequencing. The resulting plasmid was transformed into *E. coli* strain BL21-CodonPlus(DE3)-RIL. The cells were grown overnight at 37 °C in 64 ml of 2-YT media with 50 µg/ml kanamycin A and 0.4% D-glucose. The grown cell culture was transferred into 3.2 l of the same media and the cells were grown for 2–2.5 h until absorbance at 600 nm reached 0.5–0.8. 1 mM IPTG was added to the cell culture, followed by its incubation at 22 °C for 4.5 h. The cells were harvested by centrifugation at 5000 × g for 15 min at 4 °C, resuspended in 50 mM Tris-HCl, 1 mM PMSF, 5 mM 2-ME, pH 7.4 buffer, and disintegrated using a French press (IBI RAS, Pushchino, Russia). The lysate with 2 mM CaCl₂ was kept at 60 °C for 3 min, transferred into ice-water bath for 30 min, and centrifuged at 15,500 × g for 30 min at 4 °C. The supernatant was dialyzed at 4 °C: three times against 3 l of distilled water with 5 mM 2-ME, and once against 1 l of 50 mM Tris-HCl, 5 mM EDTA, 1 mM 2-ME, pH 7.4 (buffer A). The dialyzed solution was loaded onto Phenyl Sepharose 6 Fast Flow column (3 cm × 18 cm) equilibrated with buffer A. The column was flushed by 50 mM Tris-HCl, 150 mM NaCl, 1 mM 2-ME, pH 7.4 buffer. PA was eluted with 50 mM Tris-HCl, 5 mM CaCl₂, 1 mM 2-ME, pH 7.4 buffer; elution rate was 2 ml/min. The fractions with PA were collected and dialyzed at 4 °C: three times against 3 l of distilled water with 5 mM 2-ME, and once against 1 l of 10 mM sodium acetate, 1 mM 2-ME, pH 3.8 (buffer B). The dialyzed protein solution was loaded onto SP Sepharose Fast Flow column (0.9 cm × 7.5 cm) equilibrated with buffer B. PA was eluted with 10 mM Tris-HCl, 1 mM 2-ME, pH 8.0; elution rate was 1 ml/min. The fractions with PA were collected, exhaustively dialyzed at 4 °C against distilled water with 5 mM 2-ME, and stored at –70 °C. The protein yield was 6–8 mg per liter of cell culture.

The sample of recombinant wild-type (rWT) csPA was homogeneous as judged by SDS-PAGE. Protein molecular mass was determined by electrospray ionization mass spectrometry. The sample of 0.5 µM csPA in 40:60 (v/v) mixture of deionized water and acetonitrile (10 mM formic acid) was directly infused at rate of 40 µl/min into LCMS-2010EV (Shimadzu Co., Kyoto, Japan) mass detector (electrospray ionization source coupled to a single quadrupole detector). Mass spectrometry spectra were collected in positive mode from 600 to 1500 *m/z* at scan speed 125 amu/s, detector voltage 1.5 kV, and nebulizing nitrogen flow 1.5 l/min. Instrument calibration was carried out using horse heart myoglobin. The csPA sample in distilled water with 5 mM of 2-ME was preliminarily reduced by 5 mM DTT, followed by gel-filtration using Sephadex G-25 column equilibrated with distilled water. The major fraction (about 75%) of rWT csPA has a molecular mass of 11,710 Da, which is consistent with mass of the fully reduced form of TrEMBL entry E1UJ20 without initiating Met residue, lacking other posttranslational modifications. Masses of minor fractions of the sample are consistent with masses of oxidized or/and metal-loaded states of the protein.

csPA lacks Trp/Tyr residues, but UV absorption and fluorescence spectra of its sample reveal presence of Trp-containing contaminating protein(s). Their minor contribution into absorbance of rWT csPA sample at 258 nm (*D*₂₅₈) was taken into account via subtraction of the sample absorbance at 280 nm (*D*₂₈₀) multiplied by a factor of 0.625 (*D*₂₅₈ to *D*₂₈₀ ratio for water solution of L-Trp). rWT csPA concentrations were estimated spectrophotometrically using the molar extinction coefficient at 258 nm, $\epsilon_{258 \text{ nm}}$, of 2080 M^{–1} cm^{–1} and 2940 M^{–1} cm^{–1} for apo- and Ca²⁺-loaded forms of PA, respectively. The values of $\epsilon_{258 \text{ nm}}$ are derived from the molar extinction coefficient at 205 nm estimated according to [60].

2.2.3. Isolation of intact coho salmon parvalbumin

The frozen red muscles (220 g) cut from spine of coho salmon

(bought on local market and stored at –18 °C) were homogenized by blender and mixed with 400 ml of 15 mM Tris-HCl, 1 mM PMSF, 5 mM 2-ME, pH 7.5 buffer. The mixture was kept under magnetic stirring at 4 °C for 1 h, then centrifuged at 5000 × g for 30 min at 4 °C. The supernatant was subjected to addition of 2 mM CaCl₂, heating up to 60 °C (kept for 3 min), cooling using ice-water bath and centrifugation at 20,000 × g for 30 min at 4 °C. The resulting supernatant was filtered through cotton wool. The calcium-induced elution of csPA from Phenyl Sepharose 6 Fast Flow column was performed as described for rWT csPA (see above). Finally, gel-filtration using Sephacryl S-100 HR column (3 cm × 30 cm) was performed in 20 mM Tris-HCl, 150 mM NaCl, 1 mM 2-ME, pH 7.5 buffer; elution rate was 1 ml/min.

The sample of intact csPA was homogeneous as judged by SDS-PAGE. Electrospray ionization mass spectrometry measurements (see above) revealed that the major fraction (> 67%) of intact csPA has a mass of 11,752 Da, which is consistent with mass of the fully reduced TrEMBL entry E1UJ20 without Met1 residue, but containing an N-terminal acetyl group.

Intact csPA (contains Trp/Tyr-containing contaminating protein(s)) concentration was estimated spectrophotometrically using the molar extinction coefficient at 258 nm, $\epsilon_{258 \text{ nm}}$, of 1626 M^{–1} cm^{–1} and 1889 M^{–1} cm^{–1} for apo- and Ca²⁺-loaded forms of PA, respectively. The values of $\epsilon_{258 \text{ nm}}$ are derived from the molar extinction coefficient at 205 nm estimated according to [60].

2.2.4. Reduction of thiol groups in csPA

Quantitation of thiol groups of csPA (that has 4 Cys residues) was performed according to the modified Ellman's reagent test [61]. The csPA sample in distilled water with 5 mM of 2-ME was reduced by 5 mM DTT, concentrated using Amicon Ultra-4 centrifugal filters at 7000 × g for 30 min at 20 °C, and purified from 2-ME and DTT by passage through PD MidiTrap G-25 column equilibrated with 50 mM Tris-HCl, 1 mM EDTA, pH 8.0 buffer. The csPA solution (20 µM) was unfolded by 6 M GdmCl and treated by 1 mM DTNB for 10 min at 20 °C. The kinetics of the reaction was monitored by absorbance at 412 nm with a Cary 100 spectrophotometer (Varian Inc.) in a double-beam mode versus the reference solution without the protein. The molar extinction coefficient of 5-thio-2-nitrobenzoic acid at 412 nm of 13,700 M^{–1} cm^{–1} (manual to Thermo Scientific cat. #22582) was used. Efficacy of the procedure used was confirmed by the analogous examination of reduced glutathione solution. Since free thiol to rWT csPA molar ratio was of 2.8, reduction of its thiols was optimized for the highest yield of free thiols as follows. The csPA sample (0.8 mg/ml) in distilled water with 5 mM of 2-ME was treated by 5 mM DTT in 20 mM H₃BO₃-KOH, 2 mM CaCl₂, pH 8.2 buffer, followed by the solution degassing for 15 min and incubation for 1 h at 20 °C. The protein solution was concentrated by Amicon Ultra-4 filters at 7000 × g for 30 min at 20 °C and passed through PD MidiTrap G-25 column equilibrated with a buffer of choice. Free thiol to csPA molar ratio was 3.2 and 3.4 for intact and rWT PA forms, respectively, as evidenced by the modified Ellman's test. The reduced protein was used for experiments as soon as possible. In some cases, freshly prepared 0.18–1 mM DTT was added to the protein solution to avoid its oxidation.

2.2.5. Removal of metal ions from csPA

The contaminating Ca²⁺/Mg²⁺ ions in csPA samples were removed via precipitation by TCA [62] of the preliminary reduced protein (see above) in 10 mM HEPES-KOH, 1 mM DTT, pH 7.3 or 8.2 buffer, followed by the gel-filtration [63] using Sephadex G-25 column equilibrated with 10 mM HEPES-KOH, pH 7.3 or 8.2 buffer. This approach was applied for preparation of metal-free csPA for measurements of its calcium affinity (gel-filtration at pH 8.2) and limited proteolysis (gel-filtration at pH 7.3, followed by addition of 1 mM EDTA). In other cases, only the gel-filtration [63] during the final step of the csPA reduction procedure (see above) was employed using 10 mM HEPES-KOH, 1–2 mM EDTA, pH 7.3/8.2 or 10 mM H₃BO₃-KOH, 1.5 mM EDTA,

pH 8.2–9.0 or 20 mM glycine-KOH, 1 mM EDTA, pH 9.2 buffer. In either case, the resulting metal-free sample of reduced csPA is referred to as ‘apo-csPA’. The Ca^{2+} -loaded csPA was prepared in a manner identical to that for preparation of the apo-protein, but during the final gel-filtration the column was equilibrated with the analogous buffer, but using 1 mM CaCl_2 instead of 1–2 mM EDTA (in the case of limited proteolysis, 1 mM CaCl_2 was added after the gel-filtration). The respective sample of the reduced protein is referred to as ‘ Ca^{2+} -loaded’ csPA. In all cases, the final gel-filtration was carried out in the presence of freshly prepared 0.18–1 mM DTT, except for the calorimetric measurements, which are incompatible with DTT usage.

2.2.6. Circular dichroism measurements

Far-UV circular dichroism (CD) studies were carried out at 20 °C with a J-810 spectropolarimeter (JASCO, Inc.), equipped with a Peltier-controlled cell holder. The instrument was calibrated with an aqueous solution of d-10-camphorsulfonic acid and purged with nitrogen. The quartz cell with pathlength of 1 mm was used. csPA concentration was 10–16 μM . Buffer conditions: pH 8.8, 10 mM H_3BO_3 -KOH, 180 μM DTT, 1.5 mM EDTA or 1 mM CaCl_2 for apo- and Ca^{2+} -loaded csPA, respectively. The contribution of the buffer was subtracted from the experimental spectra. Band width was 2 nm, averaging time 2 s, and accumulation 3. Quantitative estimates of the secondary structure fractions were performed using CDPro software package (<http://amar.colostate.edu/~sreeram/CDPro/main.html>) [64]. The experimental data in 200–240 nm range were treated by SELCON3, CDSSTR and CONTIN algorithms, using SDP48 and SMP56 reference protein sets. The final secondary structure fractions represent the averaged values.

GdmCl-induced unfolding of apo-csPA was carried out using a set of pre-incubated for 1 h solutions with equal protein concentrations (10–13 μM), differing by GdmCl concentration. The volumetric effects accompanying dilution of GdmCl stock solution were taken into account according to [65].

2.2.7. The scanning calorimetry measurements

The differential scanning calorimetry (DSC) studies were carried out on a Nano DSC microcalorimeter (TA Instruments) at a heating rate of 1 K/min and an excess pressure of 4 bar. csPA concentration was 1.5–1.7 mg/ml. Buffer conditions: 10 mM H_3BO_3 -KOH, 1 mM CaCl_2 , pH 9.0 or 20 mM glycine-KOH, 1 mM EDTA, pH 9.2. The DSC measurements and calculations of protein specific heat capacity (C_p) were performed as described in [66]. The partial molar volume of csPA and specific heat capacity of the fully unfolded protein were estimated according to [67] and [68], respectively.

2.2.8. Limited enzymatic proteolysis of csPA

3.5 μl of bovine α -chymotrypsin solution (0.1 mg/ml) in 10 mM HEPES-KOH, 1 mM DTT, 1 mM EDTA/ CaCl_2 , pH 7.3 buffer was added to 100 μl of rWT csPA solution (0.7 mg/ml) in the same buffer. The reaction proceeded overnight at 37 °C, and terminated by addition of 1 mM PMSF in 2-propanol. The proteolytic mixture was analyzed by SDS-PAGE (15% resolving gel; 4 μg of csPA per lane). The gel was stained with Coomassie Brilliant Blue R-250 and scanned using Molecular Imager PharoFX Plus System (Bio-Rad Laboratories, Inc.).

2.2.9. Chemical crosslinking of csPA

The protein (1.2–2 mg/ml) crosslinking with 0.02% glutaric aldehyde was performed in 10 mM H_3BO_3 , 1 mM DTT, 1.5 mM EDTA or 1 mM CaCl_2 , pH 8.2 buffer. The reaction proceeded overnight at 20 °C and terminated with SDS sample buffer. The samples were subjected to SDS-PAGE (5% concentrating and 15% resolving gels; 5 μg of csPA per lane) and staining with Coomassie Brilliant Blue R-250. The gel was scanned using Molecular Imager PharoFX Plus System (Bio-Rad Laboratories, Inc.) and analyzed by Quantity One software.

2.2.10. Dynamic light scattering studies

Dynamic light scattering (DLS) measurements were carried out at 25.0 °C using a Zetasizer Nano ZS system (Malvern Instruments Ltd.). The backscattered light from a 4 mW He-Ne laser (632.8 nm) was collected at an angle of 173°. The csPA concentration was 1.5–2 mg/ml. The protein samples were preliminarily passed through 0.02 μM Whatman Anotop 10 syringe filters. The buffer conditions were pH 8.2, 10 mM H_3BO_3 , 1 mM DTT, 1.5 mM EDTA or 1 mM CaCl_2 . The acquisition time for a single autocorrelation function was 100 s. The resulting autocorrelation functions were averaged values from ten measurements. The intensity-weighted size distributions were calculated using preset parameters for distilled water at 25 °C (refractive index of 1.33 and a viscosity value of 0.8882 cP). The estimates of mean hydrodynamic radii of csPA were derived from these distributions.

2.2.11. Fluorescence measurements

Fluorescence studies were carried out at 20 °C on a Cary Eclipse spectrofluorimeter (Varian, Inc.), equipped with a Peltier-controlled cell holder. Quartz cells with pathlength of 10 mm were used. csPA and bis-ANS concentrations were 10–16 μM and 1 μM , respectively. Bis-ANS concentration was estimated spectrophotometrically using molar extinction coefficient at 385 nm, $\epsilon_{385\text{ nm}}$, of 16,790 $\text{M}^{-1}\text{cm}^{-1}$ [69]. Protein and bis-ANS fluorescence were excited at 258 nm and 385 nm, respectively. Buffer conditions: 10 mM HEPES-KOH, 0.2–1 mM DTT, pH 8.2; 1–2 mM EDTA or 1 mM CaCl_2 for apo- and Ca^{2+} -loaded csPA, respectively. The maximum intensities (I_{max}) and maximum positions (λ_{max}) of fluorescence emission spectra were obtained from fits of the spectra with log-normal curves [70] using LogNormal software (IBI RAS, Pushchino).

2.2.12. Determination of calcium affinity of csPA from the protein titrations by Ca^{2+} /EDTA

The calcium affinity of csPA at 20 °C was estimated from the spectrofluorimetric titration of the apo-protein with CaCl_2 standard solutions, followed by titration of the Ca^{2+} -loaded PA with EDTA potassium salt standard solutions. PA concentration was 10–14 μM . Buffer conditions: 10 mM HEPES-KOH, 0.2 mM DTT, pH 8.2. Intrinsic PA fluorescence intensities were corrected for the dilution effect via division of the experimental values by a factor of $(1-10^{-D_{\text{exc}}})$, where D_{exc} is the protein absorption at the excitation wavelength [71]. The experimental data were described by the *sequential metal binding scheme*, previously validated for various PAs [9,44,72,73]:



where P and M denote protein and metal ion, respectively; $K_{a,1}$ and $K_{a,2}$ are equilibrium metal association constants for the two active EF-hands of the protein. Calculation of the metal association constants was performed considering the competition between PA and EDTA for metal ions, described by the scheme [I] and the following equilibrium:



The calcium association constant for EDTA, K_{EDTA} , was calculated according to [74], using the thermodynamic constants derived from IUPAC Stability Constants Database, SC-Database v.4.79 (Academic Software, Timble, UK), and corrected for temperature and ionic strength values using van't Hoff and Davies equations, respectively. The data were globally fitted according to the schemes [I] and [II] using FluoTitr v.1.4 software (IBI RAS, Pushchino, Russia) and K_1 , K_2 and specific fluorescence intensities of the protein states as the fitting parameters. The accuracy of the K_1 and K_2 estimates was about $\pm 1/4$ orders of their magnitudes.

2.2.13. Estimation of calcium affinity of csPA using a Ca^{2+} buffer

Calcium affinity of PA at 20 °C was determined spectrophotometrically using a Ca^{2+} buffer, 1 mM EDTA in 10 mM HEPES-KOH, 1 mM DTT, pH 8.2 solution. The Ca^{2+} standard solution was added (27 μM –0.99 mM CaCl_2) to the EDTA solution for control of free Ca^{2+} concentration, $[\text{Ca}^{2+}]$, calculated using IBFluo v.1.1 software (IBI RAS, Pushchino, Russia). The effective EDTA affinity to Ca^{2+} was calculated according to [74], using the thermodynamic constants derived from IUPAC Stability Constants Database, SC-Database v.4.79 (Academic Software, Timble, UK), and corrected for temperature and ionic strength values using van't Hoff and Davies equations, respectively.

The dependence of fluorescence emission intensity of bis-ANS (1 μM), $F_{\text{bis-ANS}}$, in the presence of csPA (10 μM) on free Ca^{2+} concentration was described by the model of cooperative metal binding (P and M denote protein and metal, respectively):



where K_a is an equilibrium metal association constant, and n is a cooperativity of metal binding (Hill coefficient). The experimental data were approximated by a theoretical curve using Eq. (1):

$$F_{\text{bis-ANS}} = (F_1 + f \cdot F_2) / (1 + f), \\ f = 10^{\lg K - pM \cdot n} \quad (1)$$

The fit was achieved by variation of $\lg K$, n , F_1 and F_2 (fluorescence intensities of apo- and metal-bound PA forms, respectively) using OriginPro 9.0 software (OriginLab Corporation, Northampton, MA, USA).

2.2.14. Sequence identity calculations

The identities of amino acid sequences of two proteins (SI) were calculated using Clustal Omega (1.2.1) server (<http://www.ebi.ac.uk/Tools/msa/clustalo/> [75]) for multiple sequence alignments and Sequence Identity and Similarity (SIAS) server (<http://imed.med.ucm.es/Tools/sias.html>), considering gaps: $\text{SI} = 100 \text{ (Identical positions) / (Length of the alignment)}$.

2.2.15. Intrinsic disorder predictions

Charge-hydrophathy plots (CH-plots) were constructed as described in [76,77], using the PONDR analysis v.1 software (IBI RAS, Pushchino) [78]. The loss of charge upon N-terminal acetylation of PA was taken into account. Cumulative Distribution Function (CDF) analysis was performed as described in [76,79]. Estimates of per-residue propensity for intrinsic disorder in PAs were performed using PONDR analysis v.1 software, implementing the PONDR[®] VSL2P algorithm applicable to protein regions of any length [80]. All operations with VSL2P scores, including consideration of multiple sequences alignment (MUSCLE 3.8 algorithm [81]), averaging of the data for multiple sequences, calculations of the standard deviations and VSL2P score per residue ($< \text{VSL2P} >$) were performed using the same software.

2.2.16. Analysis of amino acid composition of PAs

The average numbers of amino acid residues and their standard deviations within PA groups were calculated using PONDR analysis v.1 software (IBI RAS, Pushchino) [78]. Calculations of amino acid compositions of separate PA sequences were performed using Geneious[®] Pro 5.6.6 Build 2012-10-15 software [82]. The Swiss-Prot database (release 2014_05 – May 14, 2014) was searched for full-length parvalbumin and oncomodulin sequences, while PA-like and putative PA sequences were excluded (the query “(name:parvalbumin OR name:oncomodulin NOT name:like NOT name:putative) AND fragment:no AND reviewed:yes” was used). Three sequences with uncertain residues (P09227, P86431 and P86432) were excluded from the found 72 entries. The sequence of northern pike α -PA (P02628) was taken from [83] (α_1 isoform).

3. Results and discussion

3.1. Isolation of coho salmon parvalbumin

Parvalbumin β -1 from coho salmon (TrEMBL entry E1UJ20; ‘csPA’) was either isolated from muscles (intact PA) or expressed in *E. coli* (recombinant wild-type, ‘rWT’, protein). Both csPA forms were purified to homogeneity as judged by SDS-PAGE (Fig. S1, panels A,C). Molecular masses of the csPA samples (11,752 Da and 11,710 Da for intact and rWT protein, respectively) determined by electrospray ionization mass spectrometry (Fig. S1, panels B,D) are consistent with mass of the fully reduced protein without Met1 residue, but with an N-terminal acetyl group (extra 42 Da) in the case of intact protein. The protein purification was based on application of hydrophobic interaction chromatography: csPA was retained by Phenyl Sepharose 6 Fast Flow medium in the absence of Ca^{2+} (5 mM EDTA, pH 7.4) and eluted with the same buffer containing CaCl_2 instead of EDTA. Hence, apo-form of csPA is characterized by high surface hydrophobicity, which decreases upon its association with Ca^{2+} . To the best of our knowledge, this principle of PA isolation is reported here for the first time.

The csPA sample was isolated and stored in the presence of 2-ME to prevent unwanted oxidation of its four Cys residues. The 2-ME molecules were cleaved from the protein by DTT treatment, followed by csPA cleaning from the reducing agents by gel-filtration (see *Materials and Methods* section for details). The resulting free thiol content was 3.2 and 3.4 thiols per csPA molecule for intact and rWT forms, respectively, as evidenced by Ellman's test. The reduced sample of csPA was purified from contaminating Ca^{2+} using combination of the protein precipitation by TCA [62] and the gel-filtration method [63]. Alternatively, only the gel-filtration during the final step of the csPA reduction procedure was employed. In either case, the resulting Ca^{2+} -free sample of reduced csPA will be referred to as ‘apo-csPA’. The Ca^{2+} -loaded sample of csPA was prepared in an identical manner, but during the final gel-filtration the column was equilibrated with the buffer containing 1 mM CaCl_2 instead of 1–2 mM EDTA (in the case of limited proteolysis, 1 mM CaCl_2 was added after the gel-filtration). Such sample of reduced csPA will be referred to as ‘ Ca^{2+} -loaded’ csPA. In all cases, except for the calorimetric measurements (incompatible with the presence of DTT), the final gel-filtration and subsequent PA studies were carried out in the presence of 0.18–1 mM DTT.

Since the experimental data for intact csPA are qualitatively identical to those for the rWT protein, for the sake of brevity we will first describe the data for rWT csPA, followed by description of differences between properties of these PA forms.

3.2. Secondary structure of rWT csPA

Analysis of the far-UV circular dichroism (CD) spectrum for Ca^{2+} -loaded csPA using CDPro software [64] gives the α -helices content of 54% (Table 1), which is similar to the value for its closest (according to BLASTp [84]) PA with known tertiary structure, northern pike β -PA (whose sequence identity, SI, with csPA is of 67%; Table 1). Calcium depletion from csPA is accompanied by a drastic decrease in its α -helicity down to 22%, while its β -sheet content is markedly elevated from 6% to 22%. Similar Ca^{2+} -induced changes in protein secondary structure were observed for pike β -PA [44]. Overall, the far-UV CD data evidence that Ca^{2+} -loaded csPA possesses a well-defined, mostly α -helical secondary structure, typical for PAs, whereas the apo-protein is mostly disordered (fraction of unordered structure is 37% versus 21% for the Ca^{2+} -loaded protein).

3.3. Conformational stability of rWT csPA

Thermal stability of csPA structure was explored by differential scanning calorimetry (DSC), which provides the most direct information on thermodynamics of protein thermal unfolding [66,85,86].

Table 1

The secondary structure fractions estimated by far-UV CD for rWT csPA, and parameters of fluorescence emission spectrum (λ_{max} , maximum position; I_{max} , maximum intensity) of hydrophobic probe bis-ANS in presence of the protein at 20 °C. csPA and bis-ANS concentrations were 10 μM and 1 μM , respectively. Buffer conditions: 10 mM $\text{H}_3\text{BO}_3\text{-KOH}$, 180 μM DTT, pH 8.8 or 10 mM HEPES-KOH, 1 mM DTT, pH 8.2 for CD and fluorescence studies, respectively; 1.5–2 mM EDTA or 1 mM CaCl_2 were added for apo- and Ca^{2+} -loaded csPA, respectively. Bis-ANS fluorescence was excited at 385 nm. The CD estimates using CDPro software [64] are compared with the literature values and the values derived from analysis of tertiary structure of Ca^{2+} -loaded northern pike β -PA (SI with csPA of 67%; X-ray PDB structure 2pvp) using DSSP algorithm [95], as implemented in PDB.

PA form	Secondary structure fractions				Bis-ANS fluorescence	
	csPA		Pike β -PA			
	α -helices, %	β -sheets, %	α -helices, %	β -sheets, %	λ_{max} , nm	I_{max} , a.u.
apo	22	22	28 ^a	n.a. ^c	505	850
Ca ²⁺	54	6	58 ^a , 56 ^b	2 ^b	507	105

^a Based on data reported in [44].

^b X-ray PDB structure 2pvp.

^c n.a., data not available.

Similarly to pike α -PA and cod PA [42], DSC data for Ca^{2+} -loaded csPA revealed a complex thermal behavior (Fig. 1a), namely, presence of two distinct heat sorption peaks with apparent mid-transition temperatures of 70 °C and 116 °C. The elevated noise level of the thermogram at temperatures above 120 °C is indicative of protein aggregation. On the contrary, the DSC data for apo-csPA in the temperature range from 10 °C to 130 °C did not exhibit measurable excess heat sorption peaks characteristic for first-order thermal transitions (Fig. 1a). Such DSC behavior is a hallmark of globally disordered proteins [66]. The absence of rigid tertiary structure in apo-csPA is supported by the significantly elevated with respect to the folded Ca^{2+} -loaded protein (from 10 °C to 50 °C) specific heat capacity of the apo-protein, reaching the level characteristic for the fully unfolded protein (Fig. 1a) [66].

Taken together, the DSC data strongly evidence the absence of rigid tertiary structure in apo-csPA. This conclusion is further supported by the data on GdmCl-induced unfolding of apo-csPA at 20 °C, monitored by its molar ellipticity at 222 nm, $[\theta]_{222\text{ nm}}$ (Fig. 1b). The gradual changes in the $[\theta]_{222\text{ nm}}$ value begin at low GdmCl concentrations, thereby implying that the protein structure is nonresistant to the denaturant action, as expected for a non-rigid protein structure. Furthermore, Ca^{2+} depletion from csPA makes it susceptible to the limited proteolysis by bovine α -chymotrypsin at 37 °C (Fig. 1c). Since bovine α -chymotrypsin is specific for aromatic residues [87], the lowered resistance of apo-csPA to chymotrypsinolysis indicates an elevated solvent exposure of Phe residues of the protein (csPA lacks Trp and Tyr), consistent with a disordering of tertiary structure in apo-csPA. The increased accessibility of hydrophobic surfaces of apo-csPA is further confirmed by fluorescence of hydrophobic probe 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid, bis-ANS [88]: maximum intensity of bis-ANS fluorescence in the presence of apo-csPA exceeds that value for the Ca^{2+} -loaded protein by a factor of 8, while bis-ANS emission spectrum maximum position (λ_{max}) for apo-csPA is blue shifted by 2 nm with respect to the λ_{max} value for the Ca^{2+} -bound protein (Table 1).

Overall, the experimental data on stability of tertiary structure of apo-csPA (Fig. 1, Table 1) consistently evidence its global disorder, in agreement with the drastic disordering of secondary structure of apo-csPA, observed by CD (Table 1). The global disorder of apo-csPA structure explains its retention on Phenyl Sepharose 6 Fast Flow column (see above). The analogous results were previously reported for metal-depleted pike PAs [9,44].

3.4. Quaternary structure of csPA

Since the increased solvent exposure of hydrophobic residues in apo-csPA could promote its multimerization, quaternary structure of csPA was probed by chemical crosslinking with 0.02% glutaric aldehyde at 20 °C, followed by SDS-PAGE (Table 2). Monomer and dimer of csPA are the predominant protein forms (total fraction of 69–88%) with bias towards the monomer (41–55%), regardless of calcium level. PA

multimerization was previously reported for intact and rWT forms of pike α -PA [9]. Overall, calcium depletion virtually does not affect multimerization state of csPA. Hence, the 43% higher with regard to the Ca^{2+} -loaded protein mean hydrodynamic radius, R_s , of apo-csPA, as evidenced by dynamic light scattering (DLS) at 25.0 °C (Table 2), reflects an expansion of the protein molecule due to its unfolding (R_s value for Ca^{2+} -loaded csPA is close to that predicted for the native monomer according to [89], see Table 2). Meanwhile, despite marked dimerization of apo-csPA, its R_s value does not reach the value predicted for the monomeric protein unfolded by GdmCl [89] (Table 2). Hence, apo-csPA remains partially structured, a conclusion supported by the CD data and the high affinity of this form to bis-ANS (Table 1).

Summing up, the combination of hydrodynamic, thermodynamic, spectral and biochemical techniques suggests the lack of defined tertiary structure in apo-csPA, which is attained upon Ca^{2+} association.

3.5. Calcium-binding properties of rWT csPA

Considering that the decreased conformational stability of apo-protein is expected to favor its higher metal affinity (discussed in [52]), Ca^{2+} affinity of csPA at 20 °C has been estimated spectrofluorimetrically using protein titrations by Ca^{2+} /EDTA and Ca^{2+} buffer experiments (Fig. 2). The titration of apo-csPA by Ca^{2+} monitored by changes in its intrinsic fluorescence intensity reveals saturation of the fluorescent signal at Ca^{2+} to protein molar ratio of two (Fig. 2a), which evidences the presence of two active Ca^{2+} -binding sites. The experimental data on the csPA titration by Ca^{2+} and EDTA can be described by the *sequential metal binding scheme* [I], earlier validated for other PAs [9,44,72,73], and competition for Ca^{2+} between the protein and EDTA (Fig. 2a). The respective equilibrium Ca^{2+} -binding constants of csPA, $K_{a,1}$ and $K_{a,2}$, are 7×10^7 1/M and 9×10^8 1/M (dissociation constants of 14 nM and 1.1 nM, respectively) (Table 3). The dependence of fluorescence emission intensity of bis-ANS in the presence of csPA (10-fold molar excess of csPA over bis-ANS) on free Ca^{2+} concentration controlled by EDTA buffer was described by the *cooperative metal binding scheme* [III] with half-transition pCa value of 8.2 and Hill coefficient of 1.7 (Fig. 2b). Thus, Ca^{2+} binding to csPA is a cooperative process with an apparent equilibrium Ca^{2+} association constant, K_a , of 1.6×10^8 1/M (dissociation constant of 6 nM), which is in line with the $K_{a,1}$ and $K_{a,2}$ values.

The overall calcium affinity of PA can be estimated from the value of total free energy change upon Ca^{2+} binding, ΔG_{Ca} , calculated as a sum of free energy changes accompanying Ca^{2+} binding to the active PA sites ($-RT \ln (K_{a,i}[\text{H}_2\text{O}])$ [90], $i = 1,2$) (Table 3). The ΔG_{Ca} value for csPA is increased compared to that for the close relative of csPA, northern pike β -PA (SI of 67%), apo-form of which is also intrinsically disordered [44]. Therefore, csPA possesses lower overall calcium affinity, despite the decreased α -helices content in its apo-form with respect to pike apo- β -PA (Table 1). Meanwhile, calcium affinity of csPA is

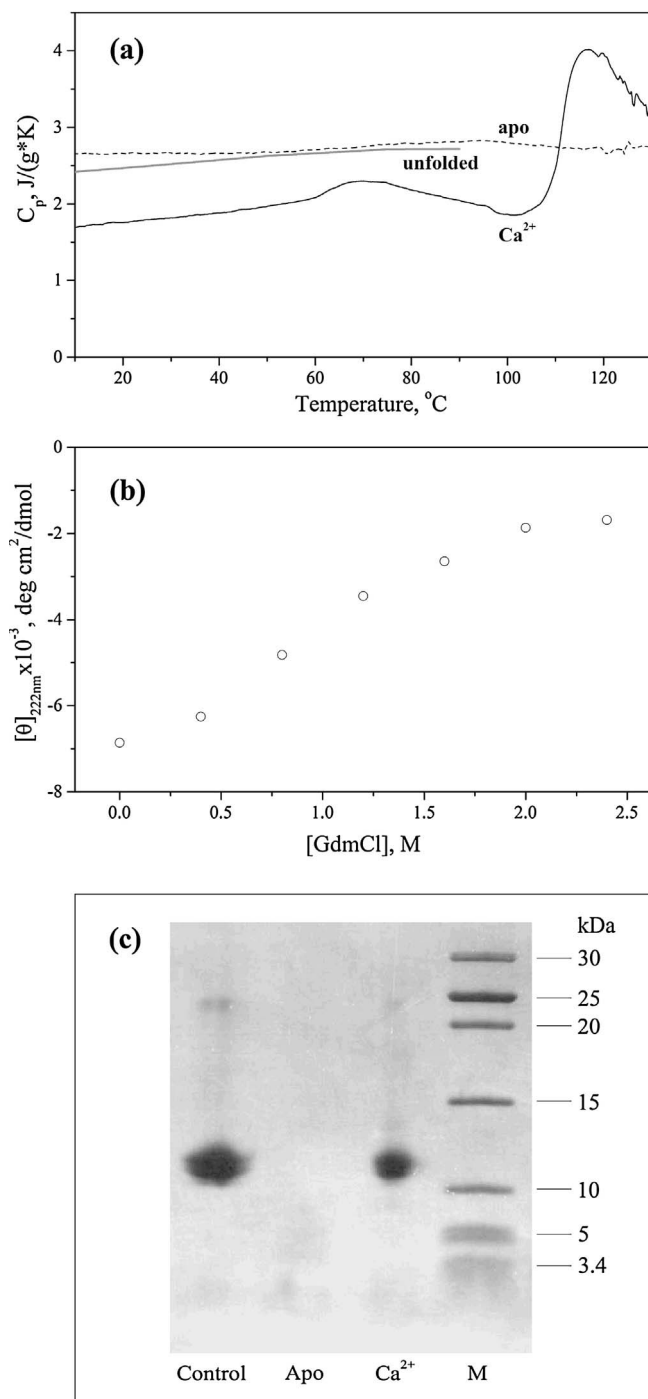


Fig. 1. Stability of tertiary structure of rWT csPA, examined by DSC (a), GdmCl-induced unfolding monitored by specific ellipticity at 222 nm (b), and chymotrypsinolysis (c). (a) Temperature dependencies of specific heat capacity of csPA, derived from DSC data. 1.5–1.7 mg/ml csPA in 10 mM H_3BO_3 -KOH, 1 mM $CaCl_2$, pH 9.0 or 20 mM glycine-KOH, 1 mM EDTA, pH 9.2 buffer, for Ca^{2+} -loaded and apo-csPA, respectively. Heating rate was 1 K/min. Grey curve corresponds to heat capacity of the fully unfolded protein, estimated according to [68]. (b) GdmCl-induced unfolding of apo-csPA at 20 °C, traced by its specific ellipticity at 222 nm. 10 μ M csPA in 10 mM H_3BO_3 -KOH, 180 μ M DTT, 1.5 mM EDTA, pH 8.8 buffer. (c) The susceptibility of csPA to overnight proteolysis by bovine α -chymotrypsin at 37 °C, monitored by SDS-PAGE (4 μ g of csPA per lane). csPA concentration was 0.7 mg/ml; the enzyme to csPA weight ratio was 1:200. Buffer conditions: pH 7.3, 10 mM HEPES-KOH, 1 mM DTT, 1 mM $CaCl_2$ or 1 mM EDTA, for Ca^{2+} -loaded and apo-csPA, respectively. The csPA sample not subjected to proteolysis was used as a control. The numbers in the right-hand column indicate molecular masses of markers ('M') in kDa.

close to that for its more distant relative, baltic cod β -PA (SI of 51%), with unfolded apo-form [53] (Table 3). Notably, apo-form of northern pike α -PA is also disordered [44], but despite higher identity of pike α -PA with csPA (SI of 59%), the calcium affinity of pike α -PA is substantially increased (Table 3). Furthermore, α isoforms of PA possess markedly higher calcium affinity with respect to the β isoforms, regardless of stability of the apo-protein: ΔG_{Ca} is -124 to -123 kJ/mol for α -PAs versus -118 to -99 kJ/mol for β -PAs (Table 3). Interestingly, several β -PAs shown to be folded at 15–25 °C possess lower calcium affinity compared to the β -PAs with disordered apo-form: ΔG_{Ca} is -112 to -99 kJ/mol versus -118 to -114 kJ/mol for β -PAs with unfolded apo-state (Table 3). Hence, the lack of stable tertiary structure in apo- β -PA seems to be associated with somewhat higher calcium affinity ($\delta \Delta G_{Ca}$ of -19 to -2 kJ/mol). Overall, while stability of apo- α -PA does not affect its calcium affinity, instability of apo- β -PA may serve as a hallmark of increased calcium affinity of the protein.

3.6. Peculiarities of physico-chemical properties of intact csPA

α -N-acetylation of PAs may affect their properties, including stability of secondary, tertiary and quaternary structure, and parameters of metal-binding [9]. To explore this possibility, we compared properties of intact csPA, which exhibits a molecular mass consistent with presence of an additional acetyl group (Fig. S1B,D), with those for the rWT protein. Apo-form of intact csPA reveals even more disorganized secondary structure (α -helices content is lowered by 4% – Table S1), more efficient penetration of bis-ANS molecules into the protein hydrophobic interior (λ_{max} value for bis-ANS is blue shifted by 14 nm – Table S1), lack of cooperative GdmCl-induced unfolding (Fig. S2), more uniform distribution over oligomeric forms, notably shifted towards a monomer (Table S2), and 4-fold higher affinity to Ca^{2+} for both EF-hands (Table S3, Fig. S3). The observed effects are much less pronounced compared to those for intact and rWT pike α -PAs [9]. Importantly, the conclusion on global disordering of csPA molecule remains valid for the intact protein.

3.7. Predictions of relative stabilities of apo-PAs

Since instability of apo-PA may serve as a factor favoring its high calcium affinity, at least in the case of β -PAs (see above), ranging of PAs by structural stability of their apo-forms could promote search for PAs possessing highest/lowest calcium affinity and, respectively, having elevated/lowest stability of their Ca^{2+} -bound forms. These features of PAs are of importance for the selection of PAs suited for use in immunotherapy of fish allergy, since PA forms with lowered Ca^{2+} affinity were previously proposed for such purposes [22,91–94], and high structural stability of food allergens is one of the factors contributing to their allergenicity (reviewed in [19]).

Considering that the two currently available binary disorder predictors (i.e., predictors evaluating the probability of the whole protein to be structured or globally disordered), CH-plot [76,77] and CDF analysis [76,79], are unable to predict instability of the apo-PAs shown to be unfolded (Table 3), some alternative approaches are required. Since low structural stability of apo-PA can be attributed to unfavorable electrostatic repulsion of carboxylates within EF-hands of the protein [44], one may expect that stability of apo-PA could be predicted from its mean charge or total negative charge of the protein molecule, or total charge of the active EF-loops. Unfortunately, examination of these opportunities for the two PA groups, differing by stabilities of their apo-forms (Table 3), did not confirm this suggestion, although there is a poorly defined tendency to decreased charge of CD-loop for PAs with disordered apo-state (-5.8 ± 1.0 versus -5.0 ± 0.7 for PAs with folded apo-form). Meanwhile, estimates of per-residue propensity for intrinsic disorder in these two PA groups using the average PONDR[®] VSL2P profiles [80] (Fig. 3) revealed a well-defined regularity: the N-terminal PA region up to residue 35 of the aligned sequence (includes

Table 2
Distribution of rWT csPA over oligomeric forms according to its chemical crosslinking with 0.02% glutaric aldehyde (20 °C), and mean hydrodynamic radius (R_s) of csPA according to DLS data (25.0 °C). 1.5–2 mg/ml csPA in pH 8.2, 10 mM H_3BO_3 , 1 mM DTT, 1.5 mM EDTA or 1 mM $CaCl_2$, for the apo- and Ca^{2+} -loaded protein, respectively. The fractions of multimeric PA forms were estimated from SDS-PAGE data (staining with Coomassie Brilliant Blue R-250). The R_s values predicted for native and unfolded forms of csPA monomer are calculated according to [89].

Protein state	Fractions of protein multimers, %					R_s , nm		
	12 kDa band	24 kDa band	36 kDa band	48–60 kDa bands	> 67 kDa band	DLS data	Prediction for a monomer	
							Native	Unfolded by GdmCl
apo	55	33	5	–	7	2.78 ± 0.04	1.77 ± 0.02	3.16 ± 0.06
Ca^{2+}	41	28	10	14	7	1.95 ± 0.03		

α -helix A, AB-loop and N-terminal half of α -helix B) is predicted to be less ordered in PAs with unfolded apo-state. Therefore, the PONDR® VSL2P score per residue for fragment 1–35 of the aligned sequence ($<VSL2P>_{1-35}$) could be used for discrimination between the folded and disordered apo-PAs. Indeed, the $<VSL2P>_{1-35}$ value of 0.609 used as a borderline between the ordered and disordered apo-PAs (criterion #1) results in an efficient discrimination between these two PA groups (Table 3). Application of the criterion #1 to 69 unambiguous PA sequences from Swiss-Prot database predicts that apo-forms for 11 of them (16%) are likely to be disordered (Table S4).
The analysis of amino acid compositions of the two PA groups with different stabilities of their apo-forms (Table 3) reveals other regularities useful for discrimination between them (Table 4). The elevated

content of Lys residues and the absence of Gln in the fragment 1–35 of the aligned sequence are hallmarks of the PAs with disordered apo-form. The increased content of Ala along with the decreased contents of Thr and Ser residues are observed in the full-length amino acid sequences of the PAs with disordered apo-state (Table 4). Meanwhile, being taken alone, none of the criteria presented in Table 4 ensures correct predictions of disorder for all apo-PAs shown in Table 3. This is achieved using the following formal criterion: apo-PA is disordered if its fragment 1–35 of the aligned sequence does not contain Gln and includes at least four Lys residues (criterion #2). Application of this criterion to the 69 unambiguous PA sequences from Swiss-Prot database shows that 13 of them (19%) are expected to be disordered in apo-state (Table S4), which is close to the estimate using the criterion #1.

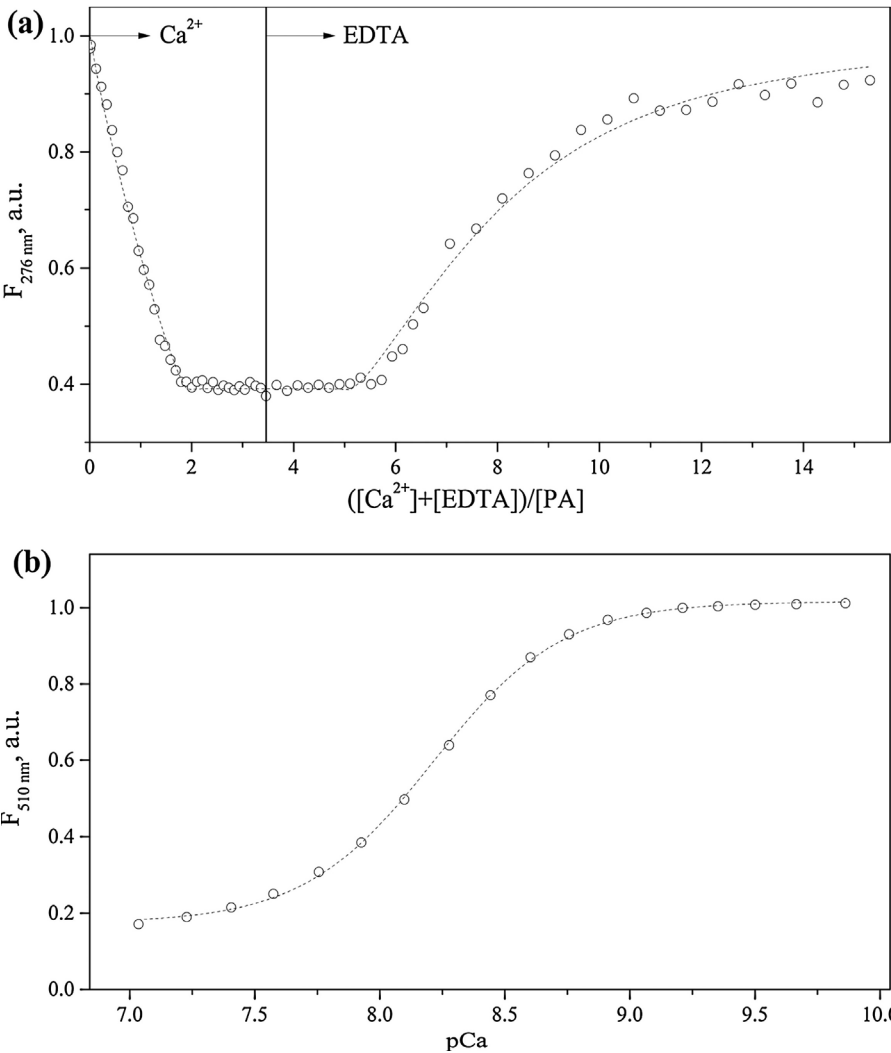


Fig. 2. Spectrofluorimetric determination of calcium affinity of rWT csPA at 20 °C from protein titrations by Ca^{2+} /EDTA (a) and Ca^{2+} buffer experiments (b). (a) The dependence of intensity of csPA (14 μ M) fluorescence (excitation at 258 nm) on total concentration of Ca^{2+} /EDTA (saturation of apo-csPA with Ca^{2+} , followed by Ca^{2+} removal by EDTA). Buffer conditions: 10 mM HEPES-KOH, 0.2 mM DTT, pH 8.2. The dashed curve is a theoretical fit to the data according to the sequential metal binding scheme [I], considering the competition for Ca^{2+} between the protein and EDTA. (b) The Ca^{2+} -dependent changes in intensity of bis-ANS (1 μ M) fluorescence (excitation at 385 nm) in the presence of csPA (10 μ M). Buffer conditions: 10 mM HEPES-KOH, 1 mM DTT, 1 mM EDTA, 27 μ M – 0.99 mM $CaCl_2$, pH 8.2. The dashed curve is a theoretical fit to the data according to the model of co-operative metal binding [III].

Table 3
The comparison of Ca²⁺ affinities of various PAs at 20 °C (pH 7.4–8.2, 10–25 mM HEPES) and experimental/predicted stabilities of their apo-forms under conditions of low ionic strength. The equilibrium Ca²⁺-binding constants of two active sites of PA, $K_{a,i}$ ($i = 1, 2$), were estimated from spectrofluorimetric titrations of PA by Ca²⁺ and Ca²⁺ chelator using the *sequential metal binding scheme* [1], or from isothermal titration calorimetry experiments as described in [96,97]. The total free energy change upon Ca²⁺ binding to PA, ΔG_{Ca} , was calculated as a sum of free energy changes accompanying Ca²⁺ binding to active PA sites ($-RT \ln (K_{a,i} [H_2O])$ [90], $i = 1, 2$). Structural stability of apo-PA was probed by various techniques, including DSC. The mid-transition temperature for thermal denaturation of apo-PA, $T_{1/2}$, was estimated from DSC data. Structural stability of apo-PAs was predicted using the CH-plots [76,77] and CDF [76,79] analyses and an original approach based on the value of VSL2P score per residue for fragment 1–35 of the aligned sequence ($< VSL2P >_{1-35}$). The borderline $< VSL2P >_{1-35}$ value of 0.609 was used for discrimination between potentially folded and unfolded apo-PAs.

Organism, PA isoform (UniProtKB entry)/intact or rWT	SI with csPA, %	$K_{a,1}, M^{-1}$	$K_{a,2}, M^{-1}$	$\Delta G_{Ca}, kJ/mol$	$T_{1/2}, ^\circ C$	References	CH-plot analysis prediction	CDF analysis prediction	$< VSL2P >_{1-35}/$ prediction
<i>Apo-form is unfolded</i>									
Northern pike, α (P02628, see α_1 in [83])/intact	59	1.2×10^9	3×10^9	−124	n/a ^a	[44]	folded	folded	0.726/unfolded
Coho salmon, β (E1UJ20)/rWT	100	7×10^7	9×10^8	−114	n/a ^a	present work	folded	folded	0.612/unfolded
Northern pike, β (P02619)/intact	67	7×10^8	5×10^8	−118	n/a ^a	[44]	folded	folded	0.768/unfolded
Baltic cod, β (P02622)/intact	51	3×10^8	3×10^8	−115	n/a ^{a,b}	[53,72]	folded	folded	0.569/folded
<i>Apo-form is folded</i>									
Rat, α (P02625)/intact	49	3×10^9 ^d	2×10^9 ^d	−123 ^d	30	[9]	folded	folded	0.560/folded
Common carp, β (P02618)/intact	70	6×10^8	2×10^7	−110	32	[54,73]	folded	folded	0.575/folded
Chicken, thymic, β (P19753)/rWT	60	2.4×10^8 ^{c,e}	1.05×10^8 ^{c,e}	−112 ^{c,e}	43.0 ^c	[56,97]	folded	folded	0.402/folded
Chicken, thymic CPV3, β (P43305)/rWT	49	4.5×10^7 ^{c,e}	2.4×10^7 ^{c,e}	−104 ^{c,e}	50.7 ^c	[57,97]	folded	folded	0.606/folded
Rat, β (P02631)/rWT	46	3.5×10^7 ^{c,e}	3.4×10^6 ^{c,e}	−99 ^{c,e}	39.8	[55,96]	unfolded	folded	0.524/folded

^a n/a, not applicable.
^b 0.1 M NaCl was used [53].
^c 0.15 M KCl [96] or 0.15 M NaCl [56,57,97] was used.
^d Temperature was 15 °C.
^e Temperature was 25 °C [96,97].

Notably, application of criteria #1 and #2 to PAs from Swiss-Prot database provide markedly different results (Table S4), which is likely related to the lower accuracy of the predictions using the criterion #1. We see no evident contradiction here, since occurrence of the unfolded apo-PAs predicted by criterion #1 is 2.5-fold higher among the apo-PAs predicted to be unfolded using the more accurate criterion #2. At the same time, comparison of the theoretical predictions using criteria #1 or #2 with the median temperature preferences of the marine fishes (TP50 [58], Table S4) does not reveal strict correlation between them. Apparently, accuracy of the both criteria should be further validated using more extensive experimental analysis.

4. Conclusions
The global disorder observed in some of apo-PAs likely contributes to their functioning as metal buffers and points out their calcium sensory role. Meanwhile, the experimentally proven cases of globally disordered apo-PAs are scarce (Table 3) and limited to cold-blooded species with a preference to cold habitat. On the contrary, the experimentally confirmed folded apo-PAs (Table 3) mostly originate from the warm-blooded species. Therefore, here we searched for PAs with disordered apo-form in a cold-blooded species, adapted to extremely cold habitat, coho salmon (TP50 of 4 °C [58]). Indeed, apo-state of recombinant csPA exhibited the whole spectrum of the

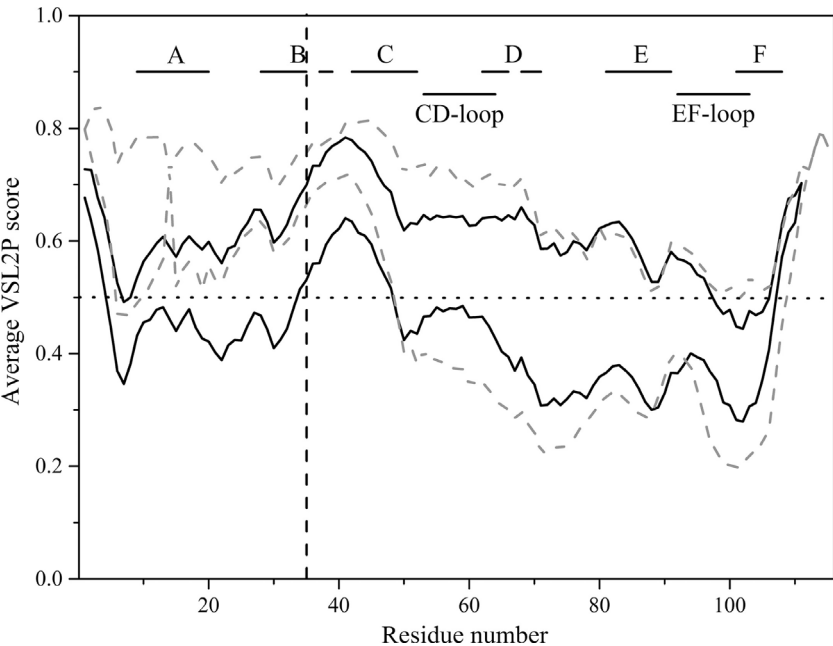


Fig. 3. Distribution over aligned sequence of the average PONDR[®] VSL2P scores [80] for the PAs with either folded (solid black curves) or unfolded (dashed grey curves) apo-states (see Table 3). The curves corresponding to highest and lowest estimates of the VSL2P scores are shown (standard deviation is either appended or subtracted from the average VSL2P curve). Positions of key secondary structure elements of PA are indicated (taken from PDB entry 2pnb): helices from A to F, and Ca²⁺-binding loops CD and EF.

Table 4

Peculiarities of amino acid compositions in the two PA groups, differing by stabilities of their apo-forms (Table 3).

Fragment 1–35 of the aligned sequence			Full PA sequence		
Residue	Number of residues		Residue	Number of residues	
	Unfolded	Folded		Unfolded	Folded
K	4.8 ± 1.3	2.4 ± 1.0	T	3.0 ± 1.2	5.0 ± 0.6
Q	0	1.6 ± 1.4	A	19.0 ± 0.7	12 ± 4
			S	4.8 ± 1.9	11 ± 3

structural features specific to globally disordered proteins, including expansion of the protein molecule, the respectively disorganized secondary structure, increased solvent exposure of hydrophobic surfaces and susceptibility to chymotrypsinolysis, elevated specific heat capacity and lack of enthalpy changes during thermal unfolding of the protein. The establishment of csPA as a new representative of the PA group with disordered apo-forms provides an opportunity to deepen our knowledge of functional peculiarities of these PAs and to uncover structural basics of their unusual behavior.

The instability of apo-csPA is expected to favor its metal affinity. Indeed, calcium affinity of csPA exceeds that of the β -PAs with experimentally confirmed folded apo-forms. In fact, the increased calcium affinity is observed for all β -PAs with unfolded apo-form. However, this is not the case for α -PAs, which could be related to very limited information on structural stabilities of apo- α -PAs. Hence, predictions of unstable apo-PAs could promote search for PAs with maximal calcium affinity, at least in the case of β -PAs.

Although electrostatic interactions are believed to contribute to destabilization of apo-PA structure, we did not reveal conclusive differences between the PA groups differing by stabilities of their apo-forms in several protein parameters, related to the intramolecular charge-charge interactions. Instead, these PA groups clearly differ in both amino acid composition and propensity for intrinsic disorder of the N-terminal fragment 1–35(36), which includes α -helix A, AB-loop and N-terminal half of α -helix B. Hence, structural stability/flexibility of AB domain is likely crucial for the maintenance of tertiary structure of apo-PA.

The two structural criteria developed for discrimination between the PAs with ordered/folded and intrinsically disordered apo-states based on the specific features of the N-terminal fragment 1–35(36) predict that 16–19% of apo-PAs are globally disordered. Since these criteria provide overlapping, but markedly different predictions, lacking clear correlation with median temperature preferences of the marine fishes, both criteria should be further validated using additional experimental analysis. The theoretical predictions presented here provide a starting point for further search of the relatively rare cases of apo-PAs with lowered structural stability. Discovery of new members of this remarkable group of PAs will enable refinement of their structural signatures.

Declarations of interest

We declare that there are no conflicts of interest.

Acknowledgement

This work was supported by grant to P.E.A. from the Russian Science Foundation (agreement #16-14-10373).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ceca.2017.08.011>.

References

- [1] J.G. Henrotte, A crystalline constituent from myogen of carp muscles, *Nature* 169 (1952) 968–969.
- [2] G. Hamoir, S. Konosu, Carp myogens of white and red muscles. General composition and isolation of low-molecular-weight components of abnormal amino acid composition, *Biochem. J.* 96 (1965) 85–97.
- [3] S. Konosu, G. Hamoir, J.F. Peckere, Carp myogens of white and red muscles. Properties and amino acid composition of the main low-molecular-weight components of white muscle, *Biochem. J.* 96 (1965) 98–112.
- [4] J.F. Peckere, Muscular parvalbumins as homologous proteins, *Comp. Biochem. Physiol.* 24 (1968) 289–295.
- [5] E.A. Permyakov, Parvalbumin, Nova Science Publishers, Inc., New York, 2006.
- [6] S.H. Arif, A Ca(2+) -binding protein with numerous roles and uses: parvalbumin in molecular biology and physiology, *Bioessays* 31 (2009) 410–421.
- [7] E.A. Permyakov, V.N. Uversky, S.E. Permyakov, Parvalbumin as a pleomorphic protein, *Curr. Protein Pept. Sci.* (2016).
- [8] C.E. Nockolds, R.H. Kretsinger, C.J. Coffee, R.A. Bradshaw, Structure of a calcium-binding carp myogen, *Proc. Natl. Acad. Sci. U. S. A.* 69 (1972) 581–584.
- [9] S.E. Permyakov, A.A. Vologzhannikova, V.I. Emelyanenko, E.L. Knyazeva, A.S. Kazakov, Y.S. Lapteva, M.E. Permyakova, A.P. Zhadan, E.A. Permyakov, The impact of alpha-N-acetylation on structural and functional status of parvalbumin, *Cell Calcium* 52 (2012) 366–376.
- [10] C.W. Heizmann, M.W. Berchtold, A.M. Rowlerson, Correlation of parvalbumin concentration with relaxation speed in mammalian muscles, *Proc. Natl. Acad. Sci. U. S. A.* 79 (1982) 7243–7247.
- [11] M. Muntener, L. Kaser, J. Weber, M.W. Berchtold, Increase of skeletal muscle relaxation speed by direct injection of parvalbumin cDNA, *Proc. Natl. Acad. Sci. U. S. A.* 92 (1995) 6504–6508.
- [12] B. Schwaller, J. Dick, G. Dhoot, S. Carroll, G. Vrbova, P. Nicotera, D. Pette, A. Wyss, H. Bluethmann, W. Hunziker, M.R. Celio, Prolonged contraction-relaxation cycle of fast-twitch muscles in parvalbumin knockout mice, *Am. J. Physiol.* 276 (1999) C395–C403.
- [13] B. Schwaller, The use of transgenic mouse models to reveal the functions of Ca(2+) buffer proteins in excitable cells, *Biochim. Biophys. Acta* (2011).
- [14] M.W. Berchtold, H. Brinkmeier, M. Muntener, Calcium ion in skeletal muscle: its crucial role for muscle function, plasticity, and disease, *Physiol. Rev.* 80 (2000) 1215–1265.
- [15] J. Zhang, V. Shettigar, G.C. Zhang, D.G. Kindell, X. Liu, J.J. Lopez, V. Yerrimuni, G.A. Davis, J.P. Davis, Engineering parvalbumin for the heart: optimizing the Mg binding properties of rat beta-parvalbumin, *Front. Physiol.* 2 (2011) 77.
- [16] W. Wang, M.S. Barnabei, M.L. Asp, F.I. Heinis, E. Arden, J. Davis, E. Braunlin, Q. Li, J.P. Davis, J.D. Potter, J.M. Metzger, Noncanonical EF-hand motif strategically delays Ca2+ buffering to enhance cardiac performance, *Nat. Med.* 19 (2013) 305–312.
- [17] M.L. Asp, F.V. Sjaastad, J.K. Siddiqui, J.P. Davis, J.M. Metzger, Effects of modified parvalbumin EF-hand motifs on cardiac myocyte contractile function, *Biophys. J.* 110 (2016) 2094–2105.
- [18] M.L. Asp, J.J. Martindale, F.I. Heinis, W. Wang, J.M. Metzger, Calcium mishandling in diastolic dysfunction: mechanisms and potential therapies, *Biochim. Biophys. Acta* 1833 (2013) 895–900.
- [19] G.A. Bannon, What makes a food protein an allergen? *Curr. Allergy Asthma Rep.* 4 (2004) 43–46.
- [20] A. Kuehn, I. Swoboda, K. Arumugam, C. Hilger, F. Hentges, Fish allergens at a glance: variable allergenicity of parvalbumins, the major fish allergens, *Front. Immunol.* 5 (2014) 179.
- [21] J.N. Stephen, M.F. Sharp, T. Ruethers, A. Taki, D.E. Campbell, A.L. Lopata, Allergenicity of bony and cartilaginous fish – molecular and immunological properties, *Clin. Exp. Allergy* 47 (2017) 300–312.
- [22] L. Zuidmeer-Jonghe, M. Fernandez-Rivas, L.K. Poulsen, A. Neubauer, J. Asturias, L. Blom, J. Boye, C. Bindslev-Jensen, M. Clausen, R. Ferrara, P. Garosi, H. Huber, B.M. Jensen, S. Koppelman, M.L. Kowalski, A. Lewandowska-Polak, B. Linhart, B. Maillere, A. Mari, A. Martinez, C.E. Mills, C. Nicoletti, D.J. Opstelten, N.G. Papadopoulos, A. Portoles, N. Rigby, E. Scala, H.J. Schnoor, S.T. Sigurdardottir, G. Stavroulakis, F. Stolz, I. Swoboda, R. Valenta, R. van den Hout, S.A. Versteeg, M. Witten, R. van Ree, FAST: towards safe and effective subcutaneous immunotherapy of persistent life-threatening food allergies, *Clin. Transl. Allergy* 2 (2012) 5.
- [23] E. Olinger, B. Schwaller, J. Loffing, P. Gailly, O. Devuyt, Parvalbumin: calcium and magnesium buffering in the distal nephron, *Nephrol. Dial. Transplant.* 27 (2012)

- 3988–3994.
- [24] B. Tong, A.J. Hornak, S.F. Maison, K.K. Ohlemiller, M.C. Liberman, D.D. Simmons, Oncomodulin, an EF-hand Ca^{2+} buffer, is critical for maintaining cochlear function in mice, *J. Neurosci.* 36 (2016) 1631–1635.
 - [25] B. Schwaller, The continuing disappearance of pure Ca^{2+} buffers, *Cell. Mol. Life Sci.* 66 (2009) 275–300.
 - [26] Y. Yin, M.T. Henzl, B. Lorber, T. Nakazawa, T.T. Thomas, F. Jiang, R. Langer, L.I. Benowitz, Oncomodulin is a macrophage-derived signal for axon regeneration in retinal ganglion cells, *Nat. Neurosci.* 9 (2006) 843–852.
 - [27] M. Smargiassi, G. Daghighi, B. Leroy, P. Legreneur, G. Toubeau, V. Bels, R. Wattiez, Chemical basis of prey recognition in thamnophine snakes: the unexpected new roles of parvalbumins, *PLoS One* 7 (2012) e39560.
 - [28] P. Steullet, J.H. Cabungcal, J. Coyle, M. Didrikson, K. Gill, A.A. Grace, T.K. Hensch, A.S. LaMantia, L. Lindemann, T.M. Maynard, U. Meyer, H. Morishita, P. O'Donnell, M. Puhl, M. Cuenod, K.Q. Do, Oxidative stress-driven parvalbumin interneuron impairment as a common mechanism in models of schizophrenia, *Mol. Psychiatry* (2017).
 - [29] G.E. Hardingham, K.Q. Do, Linking early-life NMDAR hypofunction and oxidative stress in schizophrenia pathogenesis, *Nat. Rev. Neurosci.* 17 (2016) 125–134.
 - [30] F. Filice, B. Schwaller, Parvalbumin and autism: different causes, same effect? *Oncotarget* 8 (2017) 7222–7223.
 - [31] X. Jiang, M. Lachance, E. Rossignol, Involvement of cortical fast-spiking parvalbumin-positive basket cells in epilepsy, *Prog. Brain Res.* 226 (2016) 81–126.
 - [32] I. Mahar, M.S. Albuquerque, S. Mondragon-Rodriguez, C. Cavanagh, M.A. Davoli, J.-G. Chabot, S. Williams, N. Mechawar, R. Quirion, S. Krantic, Phenotypic alterations in hippocampal NPY- and PV-expressing interneurons in a presymptomatic transgenic mouse model of Alzheimer's disease, *Front. Aging Neurosci.* 8 (2016) 327.
 - [33] A. Falco, R. Pennucci, E. Brambilla, I. de Curtis, Reduction in parvalbumin-positive interneurons and inhibitory input in the cortex of mice with experimental autoimmune encephalomyelitis, *Exp. Brain Res.* 232 (2014) 2439–2449.
 - [34] C. Konradi, E.I. Zimmerman, C.K. Yang, K.M. Lohmann, P. Gresch, H. Pantazopoulos, S. Berretta, S. Heckers, Hippocampal interneurons in bipolar disorder, *Arch. Gen. Psychiatry* 68 (2011) 340–350.
 - [35] A.Y. Wang, K.M. Lohmann, C.K. Yang, E.I. Zimmerman, H. Pantazopoulos, N. Herring, S. Berretta, S. Heckers, C. Konradi, Bipolar disorder type 1 and schizophrenia are accompanied by decreased density of parvalbumin- and somatostatin-positive interneurons in the parahippocampal region, *Acta Neuropathol.* 122 (2011) 615–626.
 - [36] M.S. Albuquerque, I. Mahar, M.A. Davoli, J.G. Chabot, N. Mechawar, R. Quirion, S. Krantic, Regional and sub-regional differences in hippocampal GABAergic neuronal vulnerability in the TgCRND8 mouse model of Alzheimer's disease, *Front. Aging Neurosci.* 7 (2015) 30.
 - [37] A. Reiner, E. Shelby, H. Wang, Z. Demarch, Y. Deng, N.H. Guley, V. Hogg, R. Roxburgh, L.J. Tippett, H.J. Waldvogel, R.L. Faull, Striatal parvalbuminergic neurons are lost in Huntington's disease: implications for dystonia, *Mov. Disord.* 28 (2013) 1691–1699.
 - [38] H. Kim, S. Ahrlund-Richter, X. Wang, K. Deisseroth, M. Carlen, Prefrontal parvalbumin neurons in control of attention, *Cell* 164 (2016) 208–218.
 - [39] B.E. Jones, Neuroscience what are cortical neurons doing during sleep? *Curr. Biol.* 26 (2016) R1147–R1150.
 - [40] S. Karunakaran, A. Chowdhury, F. Donato, C. Quairiaux, C.M. Michel, P. Caroni, PV plasticity sustained through D1/5 dopamine signaling required for long-term memory consolidation, *Nat. Neurosci.* 19 (2016) 454.
 - [41] C. Shang, Z. Liu, Z. Chen, Y. Shi, Q. Wang, S. Liu, D. Li, P. Cao, BRAIN CIRCUITS. A parvalbumin-positive excitatory visual pathway to trigger fear responses in mice, *Science* 348 (2015) 1472–1477.
 - [42] S.E. Permyakov, A.S. Karnoup, A.G. Bakunts, E.A. Permyakov, Sequence micro-heterogeneity of parvalbumin pI 5.0 of pike: a mass spectrometric study, *Biochim. Biophys. Acta* 1794 (2009) 129–136.
 - [43] E.A. Permyakov, V.N. Medvedkin, L.P. Kalinichenko, E.A. Burstein, Comparative study of physicochemical properties of two pike parvalbumins by means of their intrinsic tyrosyl and phenylalanyl fluorescence, *Arch. Biochem. Biophys.* 227 (1983) 9–20.
 - [44] S.E. Permyakov, A.G. Bakunts, A.I. Denesyuk, E.L. Knyazeva, V.N. Uversky, E.A. Permyakov, Apo-parvalbumin as an intrinsically disordered protein, *Proteins* 72 (2008) 822–836.
 - [45] V.N. Uversky, What does it mean to be natively unfolded? *Eur. J. Biochem.* 269 (2002) 2–12.
 - [46] H.J. Dyson, P.E. Wright, Coupling of folding and binding for unstructured proteins, *Curr. Opin. Struct. Biol.* 12 (2002) 54–60.
 - [47] M. Fuxreiter, I. Simon, P. Friedrich, P. Tompa, Prefolded structural elements feature in partner recognition by intrinsically unstructured proteins, *J. Mol. Biol.* 338 (2004) 1015–1026.
 - [48] J. Habchi, P. Tompa, S. Longhi, V.N. Uversky, Introducing protein intrinsic disorder, *Chem. Rev.* 114 (2014) 6561–6588.
 - [49] D. Piovesan, F. Tabaro, I. Micetic, M. Necci, F. Quaglia, C.J. Oldfield, M.C. Aspromonte, N.E. Davey, R. Davidovic, Z. Dosztanyi, A. Elofsson, A. Gasparini, A. Hatos, A.V. Kajava, L. Kalmár, E. Leonardi, T. Lazar, S. Macedo-Ribeiro, M. Macossay-Castillo, A. Meszaros, G. Minervini, N. Murvai, J. Pujols, D.B. Roche, E. Salladini, E. Schad, A. Schramm, B. Szabo, A. Santos, F. Tonello, K.D. Tsigos, N. Veljkovic, S. Ventura, W. Vranken, P. Warholm, V.N. Uversky, A.K. Dunker, S. Longhi, P. Tompa, S.C. Tosatto, DisProt 7.0: a major update of the database of disordered proteins, *Nucleic Acids Res.* 45 (2017) D219–D227.
 - [50] A. Geiger, A. Rahman, F.H. Stillinger, Molecular dynamics study of the hydration of Lennard-Jones solutes, *J. Chem. Phys.* 70 (1979) 263–276.
 - [51] N. Galamba, Water's structure around hydrophobic solutes and the iceberg model, *J. Phys. Chem. B* 117 (2013) 2153–2159.
 - [52] J.L. Gifford, M.P. Walsh, H.J. Vogel, Structures and metal-ion-binding properties of the Ca^{2+} -binding helix-loop-helix EF-hand motifs, *Biochem. J.* 405 (2007) 199–221.
 - [53] M. Laberge, W.W. Wright, K. Sudhakar, P.A. Liebman, J.M. Vanderkooi, Conformational effects of calcium release from parvalbumin: comparison of computational simulations with spectroscopic investigations, *Biochemistry* 36 (1997) 5363–5371.
 - [54] V.V. Filimonov, W. Pfeil, T.N. Tsalkova, P.L. Privalov, Thermodynamic investigations of proteins. IV. Calcium binding protein parvalbumin, *Biophys. Chem.* 8 (1978) 117–122.
 - [55] M.T. Henzl, J.D. Larson, S. Agah, Influence of monovalent cations on rat alpha- and beta-parvalbumin stabilities, *Biochemistry* 39 (2000) 5859–5867.
 - [56] A. Tan, M.T. Henzl, Evidence for a Ca^{2+} -specific conformational change in avian thymic hormone, a high-affinity beta-parvalbumin, *Biochemistry* 48 (2009) 3936–3945.
 - [57] A. Tan, L.A. Markus, M.T. Henzl, Disparate impact of the S33V mutation on conformational stability in rat beta-parvalbumin (oncomodulin) and chicken parvalbumin 3, *J. Phys. Chem. B* 114 (2010) 16171–16183.
 - [58] W.W.L. Cheung, R. Watson, D. Pauly, Signature of ocean warming in global fisheries catch, *Nature* 497 (2013) 365–368.
 - [59] H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y.M. Park, N. Buso, A.P. Cowley, R. Lopez, Analysis tool web services from the EMBL-EBI, *Nucleic Acids Res.* 41 (2013) W597–W600.
 - [60] R.K. Scopes, Measurement of protein by spectrophotometry at 205 nm, *Anal. Biochem.* 59 (1974) 277–282.
 - [61] G.L. Ellman, Tissue sulfhydryl groups, *Arch. Biochem. Biophys.* 82 (1959) 70–77.
 - [62] M. Yazawa, M. Sakuma, K. Yagi, Calmodulins from muscles of marine invertebrates, scallop and sea anemone, *J. Biochem.* 87 (1980) 1313–1320.
 - [63] H.E. Blum, P. Lehky, L. Kohler, E.A. Stein, E.H. Fischer, Comparative properties of vertebrate parvalbumins, *J. Biol. Chem.* 252 (1977) 2834–2838.
 - [64] N. Sreerama, S.Y. Venyaminov, R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis, *Anal. Biochem.* 287 (2000) 243–251.
 - [65] K. Kawahara, C. Tanford, Viscosity and density of aqueous solutions of urea and guanidine hydrochloride, *J. Biol. Chem.* 241 (1966) 3228–3232.
 - [66] S.E. Permyakov, Differential scanning microcalorimetry of intrinsically disordered proteins, *Methods Mol. Biol.* 896 (2012) 283–296.
 - [67] M. Häckel, H.J. Hinz, G.R. Hedwig, Partial molar volumes of proteins: amino acid side-chain contributions derived from the partial molar volumes of some tripeptides over the temperature range 10–90 degrees C, *Biophys. Chem.* 82 (1999) 35–50.
 - [68] M. Häckel, H.J. Hinz, G.R. Hedwig, A new set of peptide-based group heat capacities for use in protein stability calculations, *J. Mol. Biol.* 291 (1999) 197–213.
 - [69] F.J. Farris, G. Weber, C.C. Chiang, I.C. Paul, Preparation, crystalline structure and optical properties of the fluorescent probe, 4-(4-bis-1-phenylamino-8-naphthalene sulfonate), *J. Am. Chem. Soc.* 100 (1978) 4469–4474.
 - [70] E.A. Burstein, V.I. Emelyanenko, Log-normal description of fluorescence spectra of organic fluorophores, *Photochem. Photobiol.* 64 (1996) 316–320.
 - [71] E.A. Permyakov, Luminescent Spectroscopy of Proteins, CRC Press, Inc, Boca Raton, Ann Arbor, London, Tokyo, 1993.
 - [72] E.A. Permyakov, A.V. Ostrovsky, L.P. Kalinichenko, Stopped-flow kinetic studies of Ca^{2+} and Mg^{2+} dissociation in cod parvalbumin and bovine alpha-lactalbumin, *Biophys. Chem.* 28 (1987) 225–233.
 - [73] E.A. Permyakov, Parvalbumin, (2007).
 - [74] G. Schwarzenbach, H. Flaschka, Die Komplextometrische Titration, Ferdinand Enke Verlag, Stuttgart, 1965.
 - [75] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W.Z. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol.* 7 (2011).
 - [76] C.J. Oldfield, Y. Cheng, M.S. Cortese, C.J. Brown, V.N. Uversky, A.K. Dunker, Comparing and combining predictors of mostly disordered proteins, *Biochemistry* 44 (2005) 1989–2000.
 - [77] V.N. Uversky, J.R. Gillespie, A.L. Fink, Why are natively unfolded proteins unstructured under physiologic conditions? *Proteins* 41 (2000) 415–427.
 - [78] S.E. Permyakov, R.G. Ismailov, B. Xue, A.I. Denesyuk, V.N. Uversky, E.A. Permyakov, Intrinsic disorder in S100 proteins, *Mol. Biosyst.* 7 (2011) 2164–2180.
 - [79] B. Xue, C.J. Oldfield, A.K. Dunker, V.N. Uversky, CDF it all: consensus prediction of intrinsically disordered proteins based on various cumulative distribution functions, *FEBS Lett.* 583 (2009) 1469–1474.
 - [80] K. Peng, P. Radivojac, S. Vucetic, A.K. Dunker, Z. Obradovic, Length-dependent prediction of protein intrinsic disorder, *BMC Bioinf.* 7 (2006) 208.
 - [81] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.* 32 (2004) 1792–1797.
 - [82] A.J. Drummond, B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, J. Heled, M. Kearse, S. Markowitz, R. Moir, S. Stones-Havas, S. Sturrock, T.W.A. Thierry, Geneious v5.4, (2011).
 - [83] S.E. Permyakov, A.G. Bakunts, M.E. Permyakova, A.I. Denesyuk, V.N. Uversky, E.A. Permyakov, Metal-controlled interdomain cooperativity in parvalbumins, *Cell Calcium* 46 (2009) 163–175.
 - [84] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezuk, S. McGinnis, T.L. Madden, NCBI BLAST: a better web interface, *Nucleic Acids Res.* 36 (2008) W5–W9.
 - [85] P.L. Privalov, Microcalorimetry of proteins and their complexes, *Methods Mol. Biol.* 490 (2009) 1–39.

- [86] C.M. Johnson, Differential scanning calorimetry as a tool for protein folding and stability, *Arch. Biochem. Biophys.* 531 (2013) 100–109.
- [87] P. Hudaky, G. Kaslik, I. Venekei, L. Graf, The differential specificity of chymotrypsin A and B is determined by amino acid 226, *Eur. J. Biochem.* 259 (1999) 528–533.
- [88] A. Hawe, M. Sutter, W. Jiskoot, Extrinsic fluorescent dyes as tools for protein characterization, *Pharm. Res.* 25 (2008) 1487–1499.
- [89] V.N. Uversky, A.K. Dunker, Understanding protein non-folding, *Biochim. Biophys. Acta* 1804 (2010) 1231–1264.
- [90] L.D. Landau, L.P. Lifshitz, *Statistical Physics, Part 1*, 3 ed., Pergamon Press, Oxford, 1980.
- [91] A. Bugajska-Schretter, M. Grote, L. Vangelista, P. Valent, W.R. Sperr, H. Rumpold, A. Pastore, R. Reichelt, R. Valenta, S. Spitzauer, Purification, biochemical, and immunological characterisation of a major food allergen: different immunoglobulin E recognition of the apo- and calcium-bound forms of carp parvalbumin, *Gut* 46 (2000) 661–669.
- [92] I. Swoboda, A. Bugajska-Schretter, P. Verdino, W. Keller, W.R. Sperr, P. Valent, R. Valenta, S. Spitzauer, Recombinant carp parvalbumin, the major cross-reactive fish allergen: a tool for diagnosis and therapy of fish allergy, *J. Immunol.* 168 (2002) 4576–4584.
- [93] I. Swoboda, A. Bugajska-Schretter, B. Linhart, P. Verdino, W. Keller, U. Schulmeister, W.R. Sperr, P. Valent, G. Peltre, S. Quirce, N. Douladiris, N.G. Papadopoulos, R. Valenta, S. Spitzauer, A recombinant hypoallergenic parvalbumin mutant for immunotherapy of IgE-mediated fish allergy, *J. Immunol.* 178 (2007) 6290–6296.
- [94] I. Swoboda, N. Balic, C. Klug, M. Focke, M. Weber, S. Spitzauer, A. Neubauer, S. Quirce, N. Douladiris, N.G. Papadopoulos, R. Valenta, A general strategy for the generation of hypoallergenic molecules for the immunotherapy of fish allergy, *J. Allergy Clin. Immunol.* 132 (2013) 979–981 (e971).
- [95] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers* 22 (1983) 2577–2637.
- [96] M.T. Henzl, J.D. Larson, S. Agah, Influence of monovalent cation identity on parvalbumin divalent ion-binding properties, *Biochemistry* 43 (2004) 2747–2763.
- [97] M.T. Henzl, S. Agah, Divalent ion-binding properties of the two avian beta-parvalbumins, *Proteins* 62 (2006) 270–278.

Supplementary materials

Table S1. The secondary structure fractions estimated by far-UV CD for intact and rWT forms of csPA, and parameters of fluorescence emission spectrum (λ_{max} , maximum position; I_{max} , maximum intensity) of hydrophobic probe bis-ANS in presence of the protein at 20°C. csPA and bis-ANS concentrations were 10-16 μ M and 1 μ M, respectively. Buffer conditions: 10 mM H₃BO₃-KOH, 180 μ M DTT, pH 8.8 or 10 mM HEPES-KOH, 1 mM DTT, pH 8.2 for CD and fluorescence studies, respectively; 1.5-2 mM EDTA or 1 mM CaCl₂ were added for apo- and Ca²⁺-loaded csPA, respectively. Bis-ANS fluorescence was excited at 385 nm. The CD estimates were performed using CDPro software [64].

PA form	Secondary structure fractions				Bis-ANS fluorescence			
	Intact csPA		rWT csPA		Intact csPA		rWT csPA	
	α -helices, %	β -sheets, %	α -helices, %	β -sheets, %	λ_{max} , nm	I_{max} , a.u	λ_{max} , nm	I_{max} , a.u
apo	18	27	22	22	491	751	505	850
Ca ²⁺	33	17	54	6	496	73	507	105

Table S2. Distribution of intact and rWT csPA over oligomeric forms according to chemical crosslinking with 0.02% glutaric aldehyde (20°C). 1.2-2 mg/ml csPA in pH 8.2 buffer, 10 mM H₃BO₃, 1 mM DTT, 1.5 mM EDTA or 1 mM CaCl₂, for the apo- and Ca²⁺-loaded states, respectively. The fractions of multimeric PA forms were estimated from SDS-PAGE data (staining with Coomassie Brilliant Blue R-250).

csPA form	Protein state	Fractions of protein multimers, %				
		12 kDa band	24 kDa band	36 kDa band	48-60 kDa bands	>67 kDa band
Intact	apo	75	25	-	-	-
	Ca ²⁺	75	25	-	-	-
rWT	apo	55	33	5	-	7
	Ca ²⁺	41	28	10	14	7

Table S3. The equilibrium Ca^{2+} -binding constants of intact and rWT forms of csPA, $K_{a,i}$ ($i=1..2$), estimated from spectrofluorimetric titrations of PA by Ca^{2+} and EDTA using the *sequential metal binding scheme* [I], considering the competition for Ca^{2+} between the protein and EDTA. Protein concentration was 10-14 μM . Buffer conditions: 10 mM HEPES-KOH, 0.2 mM DTT, pH 8.2.

<i>csPA form</i>	$K_{a,1}, M^{-1}$	$K_{a,2}, M^{-1}$
Intact	3×10^8	4×10^9
rWT	7×10^7	9×10^8

Table S4. Theoretical predictions of global disorder in apo-PA structure for 69 unambiguous PA sequences from Swiss-Prot database (release 2014_05 - May 14, 2014) using the following criteria: 1) apo-PA is unfolded if its $\langle \text{VSL2P} \rangle_{1-36}$ value exceeds 0.609; 2) apo-PA is unfolded if its fragment 1-36 of the aligned sequence does not contain Gln and includes at least four Lys residues. The experimental data on global order/disorder in apo-PAs are referenced in Table 3. TP50 values are the median temperature preferences of the marine fishes taken from [58].

<i>Swiss-Prot entry</i>	<i>Prediction using criterion #1</i>	<i>Prediction using criterion #2</i>	<i>Experimental data</i>	<i>TP50</i>
P02616	folded	unfolded	n.a.	n.a.
P02619	unfolded	unfolded	unfolded	n.a.
P02622	folded	unfolded	unfolded	6
P02626	folded	unfolded	n.a.	n.a.
P02628	unfolded	unfolded	unfolded	n.a.
P02629	folded	unfolded	n.a.	n.a.
P05939	unfolded	unfolded	n.a.	n.a.
P30563	folded	unfolded	n.a.	n.a.
P59747	folded	unfolded	n.a.	23
P80026	folded	unfolded	n.a.	n.a.
Q90YK7	folded	unfolded	n.a.	6

Q91482	unfolded	unfolded	n.a.	9
Q9I8V0	folded	unfolded	n.a.	n.a.
O35508	folded	folded	n.a.	n.a.
P02614	folded	folded	n.a.	n.a.
P02615	folded	folded	n.a.	n.a.
P02617	folded	folded	n.a.	n.a.
P02618	folded	folded	folded	n.a.
P02620	folded	folded	n.a.	18
P02621	unfolded	folded	n.a.	11
P02623	folded	folded	n.a.	n.a.
P02624	folded	folded	n.a.	n.a.
P02625	folded	folded	folded	n.a.
P02627	folded	folded	n.a.	n.a.
P02630	folded	folded	n.a.	17
P02631	folded	folded	folded	n.a.
P05940	folded	folded	n.a.	n.a.
P05941	folded	folded	n.a.	n.a.
P0CE72	folded	folded	n.a.	n.a.
P18087	folded	folded	n.a.	n.a.
P19753	folded	folded	folded	n.a.
P20472	unfolded	folded	n.a.	n.a.
P32848	folded	folded	n.a.	n.a.
P43305	folded	folded	folded	n.a.
P51879	folded	folded	n.a.	n.a.
P56503	folded	folded	n.a.	7
P80050	folded	folded	n.a.	n.a.
P80079	folded	folded	n.a.	n.a.
P80080	unfolded	folded	n.a.	n.a.
P86741	folded	folded	n.a.	9
P86743	folded	folded	n.a.	13
P86745	folded	folded	n.a.	11
P86747	folded	folded	n.a.	11
P86748	folded	folded	n.a.	11

P86749	folded	folded	n.a.	11
P86750	folded	folded	n.a.	11
P86752	unfolded	folded	n.a.	7
P86753	folded	folded	n.a.	7
P86756	folded	folded	n.a.	17
P86757	folded	folded	n.a.	17
P86759	folded	folded	n.a.	14
P86761	folded	folded	n.a.	14
P86762	folded	folded	n.a.	14
P86764	folded	folded	n.a.	14
P86765	folded	folded	n.a.	18
P86766	folded	folded	n.a.	18
P86768	folded	folded	n.a.	n.a.
P86769	folded	folded	n.a.	n.a.
P86771	folded	folded	n.a.	26
P86773	folded	folded	n.a.	26
P86774	unfolded	folded	n.a.	15
P86775	folded	folded	n.a.	15
P86778	folded	folded	n.a.	21
P86779	folded	folded	n.a.	21
Q0VCG3	folded	folded	n.a.	n.a.
Q804W2	unfolded	folded	n.a.	n.a.
Q90YK8	folded	folded	n.a.	6
Q90YK9	folded	folded	n.a.	6
Q91483	unfolded	folded	n.a.	9

n.a., data not available

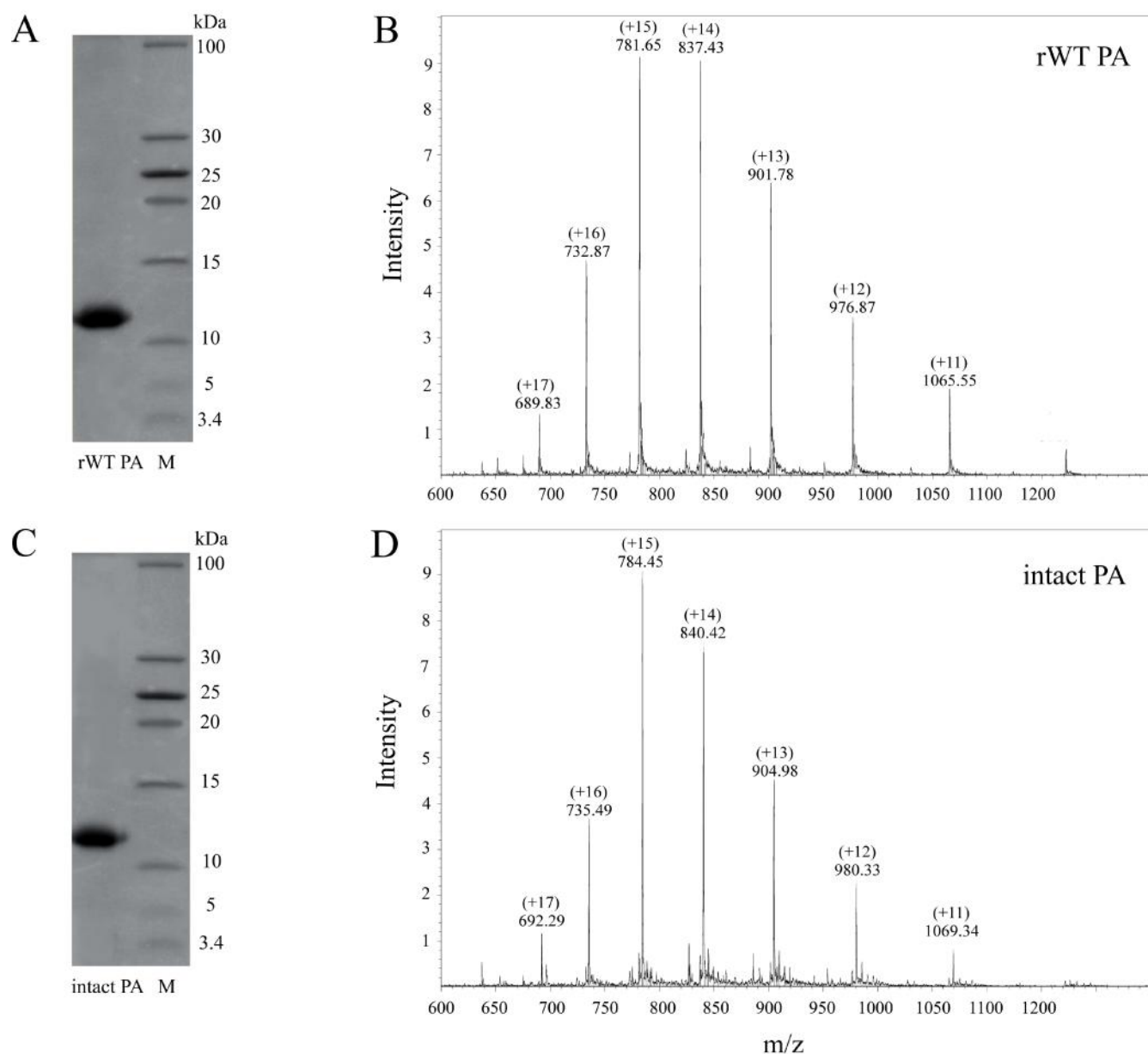


Figure S1. Reducing SDS-PAGE (15%, staining by Coomassie brilliant blue R-250; panels **A,C**) and electrospray ionization mass spectrometry (**B,D**) data for recombinant (**A,B**) and intact (**C,D**) csPA samples. ‘M’ denotes the lanes with molecular mass standards (the masses in kDa are indicated). m/z values and their respective relative charge values (+z) are indicated above the major peaks.

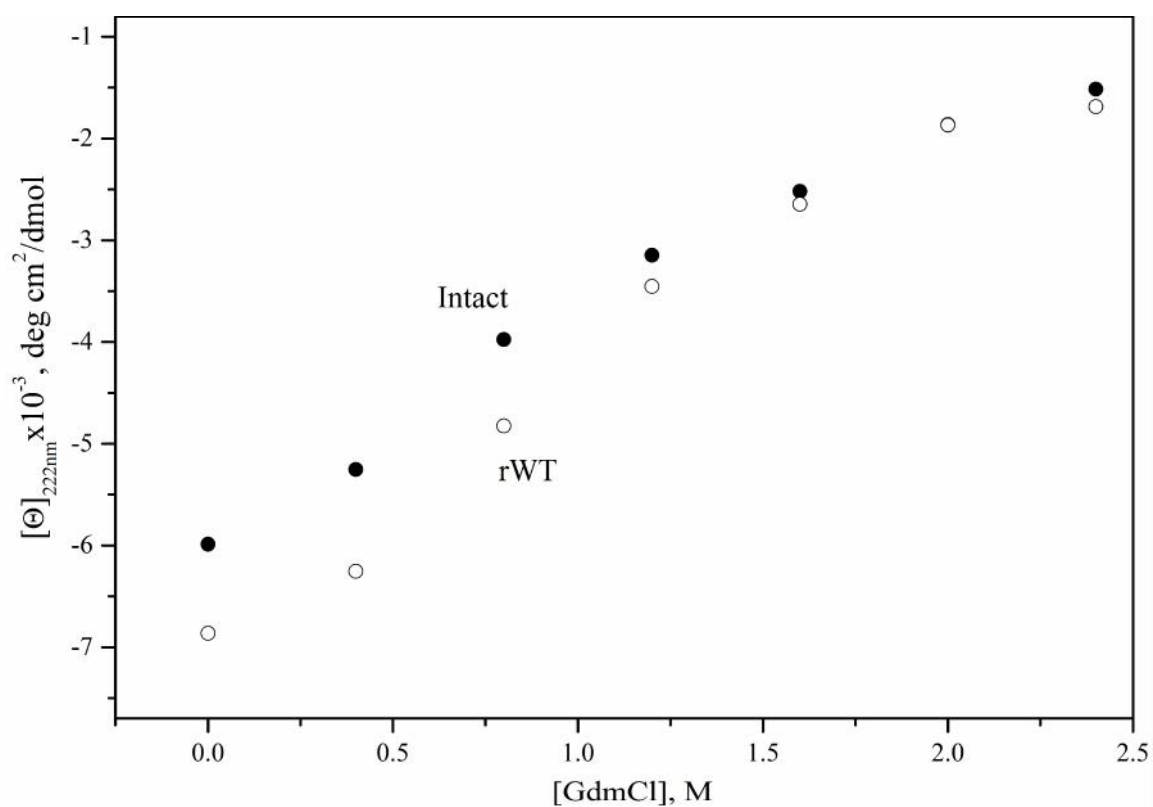


Figure S2. Stability of tertiary structure of apo-forms of intact and rWT csPA, examined by GdmCl-induced unfolding at 20°C monitored by specific ellipticity at 222 nm. 10-13 μM csPA in 10 mM H₃BO₃-KOH, 180 μM DTT, 1.5 mM EDTA, pH 8.8 buffer.

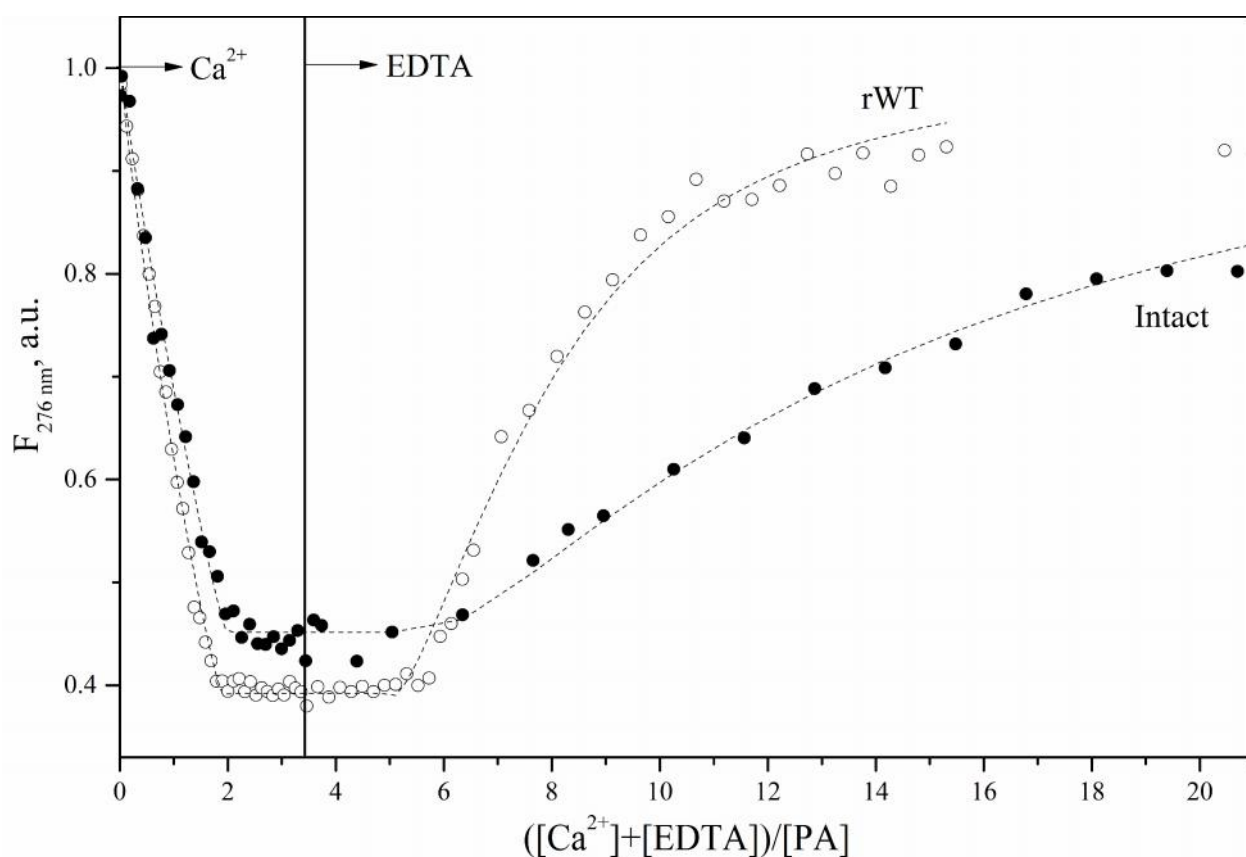


Figure S3. Spectrofluorimetric determination of calcium affinity of intact and rWT csPA at 20°C from protein titrations by Ca^{2+} /EDTA. The dependence of intensity of csPA (10-14 μM) fluorescence (excitation at 258 nm) on total concentration of Ca^{2+} /EDTA (saturation of apo-csPA with Ca^{2+} , followed by Ca^{2+} removal by EDTA). Buffer conditions: 10 mM HEPES-KOH, 0.2 mM DTT, pH 8.2. The *dashed curves* are theoretical fits to the data according to the *sequential metal binding scheme* [I], considering the competition for Ca^{2+} between the protein and EDTA.