Domain communication in Thermotoga maritima Arginine Binding Protein unraveled through protein dissection

Giovanni Smaldone, Nicole Balasco, Marilisa Vigorita, Alessia Ruggiero, Serena Cozzolino, Rita Berisio, Pompea Del Vecchio, Giuseppe Graziano, Luigi Vitagliano

PII: S0141-8130(18)32925-8
DOI: doi:10.1016/j.ijbiomac.2018.07.172
Reference: BIOMAC 10216

To appear in: *International Journal of Biological Macromolecules*

Received date: 13 June 2018
Revised date: 24 July 2018
Accepted date: 26 July 2018

Please cite this article as: Giovanni Smaldone, Nicole Balasco, Marilisa Vigorita, Alessia Ruggiero, Serena Cozzolino, Rita Berisio, Pompea Del Vecchio, Giuseppe Graziano, Luigi Vitagliano , Domain communication in Thermotoga maritima Arginine Binding Protein unraveled through protein dissection. Biomac (2018), doi:10.1016/j.ijbiomac.2018.07.172

Domain communication in *Thermotoga maritima* Arginine Binding Protein unraveled through protein dissection

Giovanni Smaldone[a,*], Nicole Balasco[b], Marilisa Vigorita[c], Alessia Ruggiero[b], Serena Cozzolino[d], Rita Berisio[b], Pompea Del Vecchio[d], Giuseppe Graziano[c], Luigi Vitagliano[b]*

[a]IRCCS SDN, Via Emanuele Gianturco 113, 80143 Napoli, Italy.

[b]Institute of Biostructures and Bioimaging, CNR, Via Mezzocannone 16, 80134 Napoli, Italy.

[c]Department of Sciences and Technologies, Università del Sannio, via Port'arsa 11, 82100, Benevento, Italy.

[d]Department of Chemical Sciences, University of Naples Federico II, via Cintia, 80126 Napoli, Italy.

*Correspondence emails: luigi.vitagliano@unina.it, gsmaldone@sdn-napoli.it

Abbreviations

TmArgBP, wild-type arginine binding protein isolated from Thermotoga maritima; $TmArgBP^{20-233}$, residues 20-233 of TmArgBP; D1, domain 1 of TmArgBP; D2, domain 2 of TmArgBP; SBP, substrate binding protein; RMSD, root mean square deviation; RMSF, root mean square fluctuation; PDB; Protein Data Bank; CD, circular dichroism; DSC, Differential scanning calorimetry; ITC, isothermal titration calorimetry; PBS, phosphate buffer saline.

Keywords: Domain-domain communication, Diagnostic tool, Protein dissection, Biosensor, Protein structure-stability, Differential scanning calorimetry.

Abstract

Substrate binding proteins represent a large protein family that plays fundamental roles in selective transportation of metabolites across membrane. The function of these proteins relies on the relative motions of their two domains. Insights into domain communication in this class of proteins have been here collected using *Thermotoga maritima* Arginine Binding Protein (TmArgBP) as model system. TmArgBP was dissected into two domains (D1 and D2) that were exhaustively characterized using a repertoire of different experimental and computational techniques. Indeed, stability, crystalline structure, ability to recognize the arginine substrate, and dynamics of the two individual domains have been here studied. Present data demonstrate that, although in the parent protein both D1 and D2 cooperate for the arginine anchoring; only D1 is intrinsically able to bind the substrate. The implications of this finding on the mechanism of arginine binding and release by TmArgBP have been discussed. Interestingly, both D1 and D2 retain the remarkable thermal/chemical stability of the parent protein. The analysis of the structural and dynamic properties of TmArgBP and of the individual domains highlights possible routes of domain communication. Finally, this study generated two interesting molecular tools, the two stable isolated domains, that could be used in future investigations.

1. Introduction

Proteins are complex biomolecules whose structure relies on the delicate balance of several distinct factors. Although they are made of thousands of atoms, minimal modifications, even a single atom, may either significantly perturb their overall structure or trigger major conformational changes. To properly elucidate the basis of protein stability/function, a deep understanding of how different regions of a protein cooperate to define its structural state is essential. Over the years, this non trivial task has been addressed by using a variety of different strategies [1, 2]. A very drastic, although highly informative, approach is to dissect a protein sequence into its main components [3]. Due to the marginal stability of the common folded state, protein dissection may have devastating consequences on its structure. Thermostable proteins are particularly suited in this field as they may allow detailed structural characterizations of the resulting fragments.

In order to understand how distant regions of a protein communicate, we used the Arginine Binding Protein isolated from *Thermotoga maritima* (TmArgBP) as model system. This protein belongs to the wide family of substrate binding proteins (SBP) [4] that cooperate with the ABC cassette system for the selective passage of metabolites through the periplasmic membrane. Extensive characterizations of this protein carried out in the last few years have highlighted that it is endowed with a number of peculiar properties. Biophysical studies have demonstrated that TmArgBP is remarkably stable against both temperature and chemical denaturants [5-10]. The crystal structure of the protein has unraveled that it forms a dimer through the swapping of the C-terminal helix **(Fig. 1)** [11], an oligomerization mechanism that frequently enables proteins with special functions [12-18].

Notably, the TmArgBP dimer combines a remarkable overall stability with a loose association of the individual subunits. TmArgBP single point mutations lead to the formation of distinct oligomeric states (monomers, dimers and trimers) [19]. Of particular interest is the observation that the trimeric forms of the protein display some amyloid-like properties, such as the ability to bind the Thioflavin T dye [19].

In this scenario, TmArgBP has also been proposed to represent a promising scaffold for the generation of an arginine sensor and an attractive model for the study of basic aspects of protein structure/function since its stability allows manipulations that are usually forbidden for proteins endowed with standard physico-chemical properties [20, 21].

We have recently analyzed the impact of the removal of the C-terminal swapping helix on TmArgBP structure/stability ($TmArgBP^{20-233}$) [22]. We here further extended this approach by dissecting the protein in the domains (D1 and D2) that constitute the bi-lobed structure of TmArgBP **(Fig. 1)**. They were characterized using a repertoire of different experimental and

computational techniques. In particular, the stability, the structure, the ability to recognize the arginine substrate, and the dynamics of the two domains have been evaluated. As for other members of the SBP family [4], the two domains cooperate for substrate binding (arginine-bound form). On the other hand, from the structural point of view, D1 and D2 are essentially independent in the ligand-free form (apo form), as they do not establish any non-covalent interaction. Our analyses provide interesting insights into TmArgBP properties. In particular, present data elucidate the role that the two domains have in substrate recognition and their mutual influence in the protein context. Moreover, our characterizations indicate that D1 may represent a promising scaffold greatly suited for Arg sensing; whereas the small size and the high thermal stability of D2 make this fragment a suitable model for accurately addressing important structure/stability relationships.

2. Materials and Methods

2.1. Protein cloning, expression and purification

The genes encoding domain 1 (D1, residues 20-114 and 207-231 linked by a GGGGSG segment) and domain 2 (D2, residues 115-206) of the protein TmArgBP (UniProt code Q9WZ62) were purchased by Sigma Aldrich. Both genes were cloned into pETM13 expression vector with a C-terminal polyHis-tag. D1 and D2 were expressed by using *Escherichia coli* BL21(DE3) cells. Induction was performed at 22°C for 16 h by the addition of 0.5 mM IPTG. Cells were then harvested and the soluble extract was obtained by sonicating cell pellets in 50 ml lysis buffer (50 mM Tris-HCl pH 8, 500 mM NaCl) in the presence of an EDTA free protease inhibitor cocktail (Roche Diagnostics). The crude cell extracts were cleared by centrifugation at 18000 rpm, and the supernatants were loaded onto a 5 ml Ni-NTA column connected to AKTA FPLC system (GE Healthcare, Milan, Italy) equilibrated with binding buffer (50 mM Tris-HCl pH 8, 500 mM NaCl). After washing with 10 volumes of binding buffer, the protein was eluted by applying a step gradient of imidazole (50-300-500 mM). To purify soluble proteins from aggregated forms or other contaminants, His-trap fractions containing D1 and D2 were loaded on Superdex S75 16/60 that was pre-equilibrated in a buffer containing Tris–HCl (pH 8.0) and NaCl 150 mM.

2.2. Notation

For comparative purposes, the starting sequence numbers of D1 and D2 were 20 and 115, respectively, to preserve the numbering scheme of TmArgBP wild-type. Due to the insertion of the GGGGSG linker, the sequence numbering of D1 C-terminal region was not coincident with that of TmArgBP. Therefore a double numbering for these residues have been reported in the text.

2.3. Light scattering

Purified proteins were analyzed by size-exclusion chromatography coupled with a triple-angle light scattering detector equipped with a QELS module (quasi-elastic light scattering). Specifically, protein samples were loaded on a S200 10/30 column, equilibrated in 50 mM Tris–HCl (pH 8.0) and 150 mM NaCl. A constant flow rate of 0.5 ml min^{-1} was applied. Elution profiles were detected by a Shodex interferometric refractometer and analyzed using a miniDawn TREOS light scattering system (Wyatt Instrument Technology Corp.). Data were processed using the Astra 5.3.4.14 software package.

2.4. CD spectroscopy

CD spectra were recorded with a J-810 spectropolarimeter equipped with a Peltier temperature control system (model PTC-423-S, Jasco Europe, Cremella (LC), Italy). Far-ultraviolet (far-UV) measurements (198–260 nm) were carried out at 20 °C using a 0.1 cm optical path length cell and proteins (15 μM) were dissolved in phosphate buffer saline (PBS) buffer (pH 7.4). CD spectra, recorded with a time constant of 4 s, a 2 nm bandwidth, and a scan rate of 10 nm min^{-1}, were signal averaged over at least three scans. The baseline was corrected by subtracting the complete buffer spectrum. Thermal denaturation curves were recorded over the 20 °C–100 °C temperature interval monitoring the CD signal at 222 nm. The curve was registered using a 0.1 cm path length cell, a protein concentration of 15 μM, and a scan rate of 1 °C min^{-1} in PBS (pH 7.4) and 4.0 M GuHCl.

2.5. Protein crystallization and X-ray data collection

Crystallization of both ligand-bound and ligand-free D1 domain and D2 domain was performed at 293 K by hanging-drop vapor diffusion methods. A screening/optimization of crystallization conditions was achieved using the crystallization solutions of Hampton Research. The best crystals of D1 domain in both the ligand-free and arginine-bound forms were obtained using a protein concentration of 6 mg ml^{-1} and 8 mg ml^{-1} respectively in a solution containing 0.2 M NaCl, 0.1M Bis-Tris (pH 5.5), 25% (w/v) polyethylene glycol (PEG) 3,350. Crystals of the D2 domain were obtained using a protein concentration of 14 mg ml^{-1} in a solution containing 0.1M Bis-Tris (pH 6.5) and 28% (w/v) polyethylene glycol 2,000.

Diffraction data for both D1 and D2 were collected in-house at 100 K using a Rigaku Micromax 007 HF generator producing Cu *Kα* radiation and equipped with a Saturn944 charge-coupled device detector. Cryoprotection of the crystals was achieved by a fast soaking in a solution containing glycerol to a final concentration of 22% (v/v). The data sets of both forms were scaled and merged using the HKL2000 program package [23]. Statistics of data collection are reported in **Table 1**.

2.6. Structure solution, refinement and validation

The crystal structures of D1 and D2 were solved by molecular replacement using the program Phaser [24]. The search fragments for the two domains were derived from the structure of the ligand-free form of the wild-type TmArgBP$^{20-247}$ (PDB code 4PRS) [11].

Crystallographic refinements were carried out using the program REFMAC [25]. A 5% of the observed data, which were randomly selected, was used in Rfree calculations to monitor the progress of refinement. Initial modelling was automatically performed using the program ARP/wARP [26]. Minor adjustments of the models were performed manually. Water molecules were incorporated into the structure in several rounds of successive refinement, followed by REFMAC runs. All residues present clear electron densities with the sole exception of the fragment 97-99 that is disordered in the crystal state of ligand-free D1. Statistics on the final models are reported in **Table 2**. The stereochemistry of the final models has been evaluated by using both standard protocols as well as innovative approaches which are based on the monitoring of the peptide geometry. In particular, we evaluated the occurrence of some correlations between geometrical parameters that are typically detected in highly accurate protein structures and are not biased by restraints in the crystallographic refinement. In particular, we checked the dependence on the ψ angle (i) of the NC$^{\alpha}$C bond angle [27], (ii) of peptide bond planarity and (iii) of the carbon carbonyl pyramidalization ϑ_C [27-32]. As shown in **Fig. S1**, the three crystallographic models display trends in the variability of NC$^{\alpha}$C bond angle and in the deviations of peptide bond from planarity that are in line with those detected in an ensemble of highly accurate protein structures [29, 30]. The atomic coordinates of D1 (ligand-free and arginine-bound forms) and D2 have been deposited in the PDB with identification codes 6GPD, 6GPC and 6GPM, respectively.

2.7. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) studies were performed at 22°C with an ITC200 calorimeter (MicroCal/GE Healthcare, Milan, Italy). A solution of L-arginine (1 mM) was titrated into a solution of D1 domain in both the ligand-free and arginine-bound forms and D2 domain (55 μM). These experiments were performed at 22°C by titrating protein solutions with drops of 2μL of L-Arginine. The analysis of the D1/D2 possible interaction was performed by titrating a solution of arginine-bound D1 domain (55μM) with drops a D2 solution at 400μM. A stirring speed of 1000 rpm was used. Fitting of data to a single-binding site model was conducted with the Origin software as supplied by GE Healthcare. ITC runs were repeated twice to evaluate the reproducibility of the results.

2.8. Differential scanning calorimetry

Differential scanning calorimetry (DSC) measurements were performed on a nano-DSC (TA Instruments) under a total pressure of 3.0 atm applied with nitrogen gas to both cells in the calorimeter. Protein samples were dialyzed against PBS solution at pH 7.4 before measurements and a scan rate of 1 °C/min was used for all DSC measurements.

The excess molar heat capacity function (<ΔCp>) was obtained after a baseline subtraction, assuming that the baseline is given by the linear temperature dependence of the native-state heat capacity. Buffer-buffer scans were recorded under the same conditions and subtracted from sample DSC profiles. The denaturation temperature (T_d) corresponds to the maximum of the DSC peak whereas the denaturation enthalpy change ($\Delta_dH(T_d)$) is the integrated area under the DSC peak. The uncertainty amounts to 0.5 °C on the T_d values and to 5% of the reported $\Delta_dH(T_d)$ numbers. The van't Hoff enthalpy change is calculated by the commonly used formula [33]:

$$\Delta_dH(T_d)_{vH} = 4R\,T_d^2\,\langle\Delta C_p(T_d)\rangle / \Delta_dH(T_d)$$

where T_d is the denaturation temperature corresponding to the maximum of DSC peak, $\langle\Delta C_p(T_d)\rangle$ is the height of the excess heat capacity at T_d, $\Delta_dH(T_d)$ is the total denaturation enthalpy change computed by direct integration of the DSC peak area, and R is the gas constant.

The finding that the calorimetric to van't Hoff enthalpy ($\Delta_dH(T_d)_{vH}$) ratio, called cooperative unit CU ($CU = \Delta_dH(T_d)/\Delta_dH(T_d)_{vH}$), is close to one is a necessary condition to state that the temperature-induced denaturation is a two-state transition [34]. The Nano-Analyze software package supplied by the manufacturer and in-house programs have been used for data analysis.

2.9. Molecular dynamics

Three independent molecular dynamics (MD) simulations were carried out on the ligand-free form of $TmArgBP^{20\text{-}233}$ [22], on the crystallographic models of D1 (ligand-free form) and D2 generated in the present study. MD simulations were performed using GROMACS software package 4.5.7 [35] with OPLS-AA as force field. The protein models were immersed in boxes filled with water molecules (TIP4P water model). Counterions were added to neutralize the systems. Equilibration was conducted in two phases. In the first phase, the temperature of the systems was stabilized at 300 K. Equilibration of pressure at 1 atm was then conducted under an NPT ensemble. Energies were first minimized by fixing the protein atoms to allow relaxation of the solvent and successively without restraints.

Before starting constant temperature MD at 300 K, 100 ps runs were performed at 50, 100, 150, 200, 250, and 300 K temperature. The timescale of the simulations was 100 ns, with a time step of 2 fs. Bond lengths were constrained according to the LINCS algorithm. Electrostatic interactions were treated using the Particle Mesh Ewald (PME) method (grid spacing of 0.12 nm). A no-bonded cutoff of 10 Å was used to account for Lennard-Jones interactions.

To monitor whether the simulation reached an adequate convergence in the essential space, we calculated the root mean square inner product (RMSIP) between the two halves of the equilibrated trajectory [36, 37]. In detail, the equilibrated region of the three trajectories (10-100 ns) was divided in two halves (10-55 ns and 55-100 ns). The values of the RMSIP (0.85 for TmArgBP$^{20-233}$, 0.77 for D1, and 0.77 for D2) between these two trajectory portions indicate that all three simulations reached a satisfactory level of convergence. The quality of the simulations was also evaluated by analyzing trajectory structures using GROMACS routines and the program VMD [38].

3. Results

3.1. TmArgBP domain design and production

Domain design was performed by considering the crystallographic structure of TmArgBP [11] and the recent independent study carried out on the histidine binding protein HisJ [3]. To avoid complicacies related to the dimeric wild-type protein (domain swapping and some heterogeneity in the oligomeric state), domain design was realized on the monomeric truncated form of the protein TmArgBP$^{20-233}$, in which the swapping C-terminal helix has been deleted [22].

The analysis of the bi-lobed structure of the protein provides straightforward hints for TmArgBP dissection. As shown in **Fig. 1**, D2 is comprised of a continuous stretch of the polypeptide chain (residues 115-206), whereas D1 is made up of two fragments (residues 20-114 and 207-231). In the final construct of the individual D1 these two fragments have been linked by means of a Gly-rich junction (GGGGSG). The sequences of D1 and D2 are reported in **Fig. 1**.

3.2. D1 and D2 biophysical characterization

Both D1 and D2 have been expressed in *E. coli* in high yields and very good homogeneity, as demonstrated by gel filtration chromatography and SDS-gel (data not shown). Light scattering analyses showed that both domains are monomeric (**Fig. 2**). CD spectra indicate that D1 and D2 are well folded (**Fig. 3**). In particular, although both spectra are indicative of proteins endowed with an alpha-beta structure, CD data suggest that D2 presents a larger content of β-structure. This observation is in line with the secondary structure content determined on the basis of TmArgBP X-ray structure [11]: the β-structure percentage of D1 and D2 is 23.2 and 26.1, respectively.

Thermal stability of the two domains was preliminary evaluated by recording the CD signal at 222 nm while increasing the temperature. As for TmArgBP and its truncated variant TmArgBP$^{20-233}$, both D1 and D2 do not undergo structural transitions in the 20-100 °C temperature range (**Fig. S2**). To gain insight into the stability of these domains, we monitored the thermal variation of the CD signal at 222 nm in the presence of 4.0 M GuHCl, as previously reported for TmArgBP [11]. Structural transitions were observed for both systems as D1 and D2 present denaturation temperatures of 58 and 71 °C, respectively (**Fig. 4)**. It is worth mentioning that the transition of D2 is fully reversible, whereas the recovering of secondary structure upon cooling D1 is not complete. In both cases, the unfolded state at 100 °C retains a significant level of secondary structure (**Fig. S3**).

The D2 stability is comparable to that of TmArgBP$^{20-233}$ which displays a denaturation temperature of 70 °C in 4.0 M GuHCl. On the other hand, extraction-isolation of D1 from the protein context leads to a significant destabilization of its folded structure.

3.3. Analysis of D1 and D2 (putative) mutual interactions

Once obtained the individual domains in a soluble and folded state, we evaluated their (putative) mutual affinity by gel filtration analyses and isothermal titration calorimetry (ITC) experiments. These studies were conducted either in the presence or absence of Arg, which is the natural ligand of TmArgBP. Gel filtration analyses did not indicate the formation of larger molecular species upon mixing the two domains (data not shown). This finding was confirmed by ITC experiments that show that the two domains do not interact (**Fig. 5**). This finding suggests that the covalent linker joining the domains is essential for the correct assembly of the wild-type protein.

3.4. Arginine recognition by D1 and D2

The purified domains were also used to check their individual ability in Arg binding/recognition. It is worth mentioning that, as shown in **Fig. 1,** the two domains cooperate in Arg anchoring by TmArgBP, as the substrate makes direct interactions with residues belonging to both D1 and D2. ITC studies were performed at 22°C by titrating L-arginine into solutions containing one of the domains. Since freshly prepared TmArgBP and TmArgBP$^{20-233}$ samples present an Arg molecule in their binding pockets, D1 and D2 were extensively dialyzed against 50mM Tris-HCl (pH 8) 150 mM NaCl to remove any ligand potentially bound during their preparation. As shown in **Fig. 6**, the two domains have distinct behavior in Arg recognition. D1 binds Arg with a remarkable affinity: at 22°C, the binding constant K_b = 3.6 (±0.4)·10^6 M^{-1}, the binding enthalpy Δ_bH = -54 (±2) kJ mol^{-1}, and the binding entropy $T{\cdot}\Delta_bS$ = -17 (±2) kJ mol^{-1}. These numbers indicate that Arg association to D1 is a highly exothermic process, in line with the several H-bonds in which Arg is involved in the

binding pocket. On the other hand, D2 is unable to anchor Arg, despite the role that this domain plays in substrate recognition by the wild-type TmArgBP (see **Fig. 1**). These results are in line with those obtained for the two domains of HisJ, even though the K_b constant for His binding to D1 was smaller by two orders of magnitude (i.e., $K_b = 1.3 \cdot 10^4$ M^{-1} at room temperature), and was detected only by means of NMR measurements [3]. The quantitative difference may be a consequence of the different flexibility-rigidity of the side chains lying on the binding pocket of the mesophilic HisJ D1 with respect to those in the hyperthermophilic TmArgBP D1.

3.5. Crystal structures of D1 and D2

To gain insights into the structural properties of TmArgBP individual domains and to unravel the impact of the context on their structure/stability, we undertook crystallographic studies on both D1 and D2. Crystals suitable for X-ray diffraction analyses were grown for both domains using PEG as precipitating agent (see Methods for details). High resolution data were collected for all these crystals (**Table 1**). The three refined models display good values for the crystallographic indicators and excellent stereochemistry (see **Table 2**).

The D1 structure indicates that there are only overall marginal variations when compared to its structure in the protein context (**Fig. 7**). The root mean square deviation (RMSD) value of the ligand-free D1 *versus* the same domain inserted in the ligand-free TmArgBP$^{20\text{-}233}$ is 0.27 Å. A slightly higher RMSD value is observed when the arginine-bound forms are compared (0.38 Å). As shown in **Fig. 8**, the arginine binding to D1 has minimal structural effects (**Fig. 8**). Rather low RMSD values are also observed when the D2 structure is compared to that of the domain in TmArgBP$^{20\text{-}233}$ (0.30 Å). Comparison of Arg anchoring by D1 and TmArgBP$^{20\text{-}233}$ unveils only subtle differences (**Fig. S4**). The D2 structure shows that the residues contributing to Arg recognition are fully exposed to the solvent (**Fig. 9**). This suggests that the inability of D2 to anchor Arg is not due to binding site rearrangements in the isolated domain, but rather to the intrinsic weakness of the specific D2-Arg interactions. The hinge region connecting the two domains assumes a well-defined β-structure in wild-type dimeric TmArgBP as well as in monomeric TmArgBP$^{20\text{-}233}$ [11, 22]. It has been proposed that the structure of the hinge region is important for the modulation of inter-domain distance in TmArgBP [22]. This hinge region was broken to generate D1 and D2 (**Fig. 1**). However, despite truncation, the dangling segments of the hinge region show an intrinsic tendency to form a β-structure in both domains (**Fig. 7**). In D1, H-bonds are formed by the dipeptide Phe112-Asp113 with the dipeptide Gln122-Tyr123 (Gln208-Tyr209 in the wild-type enzyme) whereas in D2 the N-terminal end 116-QVIVVR-121 associates with the main body of the protein by forming a β-strand.

3.6. Thermodynamic characterization

The thermal stability of D1 and D2 has been characterized by means of DSC measurements at pH 7.4 in PBS buffer, and compared to that of the parent monomeric form TmArgBP$^{20\text{-}233}$. Both the ligand-free and arginine-bound forms of the latter show large DSC peaks with T_d = 108.5 and 105.3 °C and $\Delta_dH(T_d)$ = 1100 and 1000 kJ mol^{-1}, respectively (**Fig. 10A**). The high T_d values are in line with the hyperthermophilic origin of TmArgBP$^{20\text{-}233}$ and indicate that this protein is very stable against temperature, regardless of the dimeric structure loss. A rationalization of the fact that the ligand-free form of TmArgBP$^{20\text{-}233}$ is more stable than the arginine-bound form against temperature has already been provided [22].

To evaluate the reversibility of the temperature-induced denaturation of TmArgBP$^{20\text{-}233}$, a second heating, upon cooling the sample, was performed on both the ligand-free and arginine-bound forms of the protein. This experiment does not produce a detectable endothermic peak, indicating that the thermal denaturation of both the ligand-free and arginine-bound forms of TmArgBP$^{20\text{-}233}$ is an irreversible process. However, by stopping the first heating at the T_d value, the second heating is characterized by a DSC peak (**Fig. 10B**) whose area amounts to 75% of that of the complete scan (**Fig. 10A**). This suggests that the irreversibility observed for TmArgBP$^{20\text{-}233}$ is likely due to side-reactions occurring at very high temperature [39, 40] as the temperature scan ends at 120 °C, and the protein sample spends 40 minutes above 100 °C. Therefore, the irreversibility of the process should not be an intrinsic feature of the thermal denaturation of the protein [40, 41]. The latter can be considered reversible and well described by the two-state N ⇔ D transition model, because the cooperative unit (CU) value is close to one, a necessary condition for the occurrence of a two-state transition (**Table 3**) [33, 34].

Also in the case of D1, DSC measurements have been performed on both the ligand-free and arginine-bound forms, as ITC measurements have shown that D1 is able to bind Arg at room temperature (**Fig. 6**). However, since the binding proves to be strongly exothermic (Δ_bH = -54 (±2) kJ mol^{-1}), the equilibrium binding constant decreases markedly with temperature, becoming not so large around 90 °C. In addition, on raising the temperature, the side chains of the binding pocket should gain conformational freedom weakening the attractive interactions with Arg.

This consideration implies that the arginine-bound form of D1 loses Arg on increasing the temperature. Therefore, there is no difference in the DSC traces produced by the two forms of D1. The DSC peaks of both the ligand-free and arginine-bound forms of D1 (**Fig. 11A)** are characterized by T_d= 95.5 °C and $\Delta_dH(T_d)$ = 450 kJ mol^{-1} (**Table 3**). The DSC trace of D2 domain (only ligand-free form because D2 is not able to bind Arg) (**Fig. 12**A) is characterized by T_d =

102.6 °C and $\Delta_dH(T_d)$ = 470 kJ mol^{-1}. Even though D1 is very stable against temperature, its T_d value is 10 °C lower than that of the ligand-free form of TmArgBP$^{20-233}$. On the other hand, D2 proves to have a higher stability against temperature that is closer to that of the ligand-free form of the parent protein TmArgBP$^{20-233}$.

According to the re-heating criterion, the reversibility of the temperature-induced denaturation of the two domains proved to be around 50%. However, by stopping the first heating at the respective T_d values for both D1 and D2, the second heating of the same sample produced a DSC peak exactly superimposable to that obtained in the first heating of another protein sample (**Figs 11B** and **12B)**. This can be considered a demonstration that the process is reversible for both domains and is well described by the two-state N $\Leftrightarrow$ D transition model as emphasized by the closeness to one of the CU value.

A comparison among the DSC traces of the three proteins is reported in **Fig. 13**. The marked difference in the DSC peak height does not imply a difference in cooperativity of the temperature-induced denaturation as the DSC peak height simply rises on increasing the $\Delta_dH(T_d)$ value. However, the recorded increase in $\Delta_dH(T_d)$, 1100 kJ mol^{-1} for the ligand-free form of TmArgBP$^{20-233}$ versus 450 and 470 kJ mol^{-1} for D1 and D2 domains (**Table 3**), respectively, indicates unequivocally that the hinge region is able to strongly couple the two domains. It is worth mentioning that in the ligand-free form of TmArgBP$^{20-233}$, the two domains do not form direct interactions [22, 42, 43].

Due to the non-perfect reproducibility of DSC baseline at high temperatures (also for the non-perfect reversibility of the temperature-induced denaturation), it is very difficult to obtain reliable values for the denaturation heat capacity change, Δ_dC_p, directly from DSC traces. On the basis of a large set of experimental data, Rees and Robertson proposed a simple relationship, $\Delta_dC_p = 0.058 \cdot N_{res}$, that allows the Δ_dC_p calculation from the knowledge of the number of residues of the protein (in the assumption that it is a temperature-independent quantity) [44].

Using this relationship, Δ_dC_p (in kJ K^{-1}mol^{-1}) = 12.3 for TmArgBP$^{20-233}$, 7.3 for D1, and 5.3 for D2. The knowledge of T_d, $\Delta_dH(T_d)$ and Δ_dC_p can be exploited to calculate, via the Gibbs-Helmholtz equation, the denaturation Gibbs energy change, Δ_dG, as a function of temperature. At 25 °C, Δ_dG = 119 kJ mol^{-1} for the ligand-free form of TmArgBP$^{20-233}$, 33 kJ mol^{-1} for the ligand-free form of D1, and 51 kJ mol^{-1} for D2. Normalization to the number of residues leads to a stabilization Gibbs energy of 560, 260 and 530 J mol res^{-1} for TmArgBP$^{20-233}$, D1, and D2, respectively. These numbers indicate that TmArgBP$^{20-233}$ and D2 are very stable proteins, whereas the stability of D1 is

normal (see Table 1 in Rees and Robertson to gain a right perspective [44]). However, the significance of these values should not be overemphasized in view of the long temperature extrapolation performed; since all the three proteins have T_d values around 100 °C.

3.7. Molecular dynamics studies

MD simulations have been performed to gain insights into the role that protein dynamics play in the stability of the ligand-free form of TmArgBP$^{20\text{-}233}$ and its individual domains. The structural stability of the systems in the trajectory timescale (100 ns) was evaluated by computing standard indicators, such as root mean square deviations (RMSD) from the starting model (**Fig. 14**), and secondary structure evolution (**Fig. S5**).

The time evolution of the protein secondary structure elements (**Fig. S5A**) and of the gyration radius (data not shown) indicates that TmArgBP$^{20\text{-}233}$ retains its fold throughout the simulation. Nevertheless, the large fluctuation of the RMSD values of TmArgBP$^{20\text{-}233}$ indicates that the protein is highly flexible (**Fig. 14A**). The comparison of the overall RMSD values with those displayed by the two domains in the protein context shows that mutual domain reorientations represent the main component of TmArgBP$^{20\text{-}233}$ motions. This observation is corroborated by the analysis of the distances between the centers of mass of D1 and D2 (**Fig. 14D**). This behavior of TmArgBP$^{20\text{-}233}$ is somewhat expected, taking into account the looseness of the protein in the ligand-free form, and is in line with the functional Venus fly-trap mechanism of TmArgBP and with similar studies carried out on other SBP proteins [3]. As the hinge region linking the two domains (residues 111-116 and residues 204-209) is the only physical connection between D1 and D2, we evaluated its dynamic properties in the simulation. Interestingly, although its secondary structure content is transient (**Fig. S5A**), its mobility is limited by a number of local H-bonding interactions that take place in the trajectory structures (**Fig. S6**). It is also worth mentioning that some of these residues display a rather narrow distribution in the Ramachandran plot (**Fig. S7**).

The analysis of the MD simulations on the two domains indicates that both of them retain their folded state during the trajectory (**Figs 14B,C and S5B,C**). Inspection of the root mean square fluctuations (RMSF) of the individual domains shows that D1 is more flexible than D2 (**Figs 15**). The average RMSF values are 0.82 Å for D1 and 0.61 Å for D2. When the domain mobility is monitored in the TmArgBP$^{20\text{-}233}$ context, small but significant differences emerge (**Fig. 15**). In this situation, the average RMSF values are 0.70 Å for D1 and 0.66 Å for D2. This finding implies that, even though the two domains interact only through the connecting hinge region, a mutual influence (i.e., correlation) of their motions is observed. The presence of nearby more rigid D2 makes D1 more rigid and, conversely, the D1 flexibility is somewhat transmitted to D2. In this respect, it should be underscored that the relationship between flexibility-rigidity and thermal stability of

globular proteins is not a trivial one. Actually, starting from different experimental and theoretical routes, it has been proposed that a larger flexibility of the native state could be the molecular feature of hyperthermophilic proteins to cope with high temperature (i.e., the extra thermal stability would have an entropic origin) [45-47].

4. Discussion

As for the other members of the substrate binding protein family, the biological activity of TmArgBP strongly relies on the dynamic properties of the two domains that constitute the bi-lobed structure of the protein. In SBP proteins the concerted motions of the two domains cooperate to drive substrate binding and release according to the so-called Venus fly-trap mechanism [48]. Although hundredths of distinct members of this family have been identified and characterized, some aspects of domain-domain communication deserve further investigations.

Following the approach independently adopted for HisJ [3], we have characterized the properties of TmArgBP individual domains (**Fig. 16**). By using a repertoire of different experimental and theoretical techniques, we evaluated the stability, the structure, the ability to recognize TmArgBP substrate, and the dynamics of the two domains. These analyses were performed using the truncated monomeric form of the protein ($TmArgBP^{20\text{-}233}$) as comparative model [22]. In particular, to highlight subtle domain-domain communication, most of the comparisons were conducted using the ligand-free variant of the protein as in this form the two domains are simply connected by a double strand hinge region, without making any other interaction.

Our data indicate that, although in the parent protein both D1 and D2 cooperate for Arg anchoring, only D1 is able to bind the substrate. D1 retains a significant affinity for Arg (i.e., $K_b = 3.6 \cdot 10^6\ M^{-1}$ at 22 °C), even though it is ten-fold lower than that of $TmArgBP^{20\text{-}233}$ [22]. D2 inability to recognize the substrate opens the question about the local structure of D2 residues that are deputed to Arg binding in the native protein. The D2 crystal structure shows that these D2 residues are fully exposed to the solvent and therefore, potentially, prone to Arg binding. Therefore, D2 inability to bind Arg should be due to the intrinsic weakness of the interactions that it forms within the substrate. The distinct role that the two domains have in substrate recognition is in line with the recent results reported for *E. coli* HisJ [3]. Collectively these findings suggest that this is likely a general feature of the vast class of SBPs. The D1 ability to retain a significant affinity for Arg has some interesting implications. This suggests that the role of the ABC cassette receptor is not limited to favor the opening of the SBP partner, but it should also induce conformational changes that reduce the affinity of the substrate for D1. Although present binding assays corroborate this concept, caution should be used in their extrapolation to the physiological conditions of

Thermotoga maritima. Since data indicate that Arg binding to D1 is highly exothermic (i.e., ΔH = -54 kJ mol^{-1} at room temperature), its strength decreases strongly on increasing the temperature. In this scenario, minimal modifications of D1 structure may allow the substrate release.

Measurements indicate that D1 and D2 retain the remarkable thermal and chemical stability of the parent $TmArgBP^{20-233}$ despite fragmentation. Taking into account the different stability of the two domains and the absence of interactions among them, except the covalent linkage (**Fig. 1**), it is somewhat surprising that a single DSC peak is observed in the temperature-induced denaturation of the parent protein. This observation indicates that the hinge region joining D1 and D2 is able to transfer information from one domain to the other, coupling their motions in an efficient and robust manner. As emphasized by the analysis of MD trajectories, the flexibility of each domain is significantly affected by the presence of the other.

In conclusion, the present study has provided not only interesting clues on basic aspects of protein stability, but also has generated interesting molecular tools that could be used in further investigations. Arginine sensing is an important issue in different fields that include the diagnosis of severe pathologies and the food quality control [49, 50]. The reduced size, the remarkable stability and the relevant Arg affinity make D1 a promising scaffold for the development of an Arg sensor. Moreover, the superficial binding of the ligand also suggests that D1 variants aimed at sensing other amino acids may be designed. Finally, the small size and the denaturation reversibility of D2 make this domain a suitable model for accurately addressing structure/stability relationships.

References

[1] W.S. Bennett, R. Huber, Structural and functional aspects of domain motions in proteins, CRC Crit Rev Biochem 15(4) (1984) 291-384.
[2] B.H. Oh, C.H. Kang, H. De Bondt, S.H. Kim, K. Nikaido, A.K. Joshi, G.F. Ames, The bacterial periplasmic histidine-binding protein. structure/function analysis of the ligand-binding site and comparison with related proteins, J Biol Chem 269(6) (1994) 4135-43.
[3] B.C. Chu, T. DeWolf, H.J. Vogel, Role of the two structural domains from the periplasmic Escherichia coli histidine-binding protein HisJ, J Biol Chem 288(44) (2013) 31409-22.
[4] R.P. Berntsson, S.H. Smits, L. Schmitt, D.J. Slotboom, B. Poolman, A structural classification of substrate-binding proteins, FEBS Lett 584(12) (2010) 2606-17.
[5] M.S. Luchansky, B.S. Der, S. D'Auria, G. Pocsfalvi, L. Iozzino, D. Marasco, J.D. Dattelbaum, Amino acid transport in thermophiles: characterization of an arginine-binding protein in Thermotoga maritima, Mol Biosyst 6(1) (2010) 142-51.
[6] A. Scire, A. Marabotti, M. Staiano, L. Iozzino, M.S. Luchansky, B.S. Der, J.D. Dattelbaum, F. Tanfani, S. D'Auria, Amino acid transport in thermophiles: characterization of an arginine-binding protein in Thermotoga maritima. 2. Molecular organization and structural stability, Mol Biosyst 6(4) (2010) 687-98.
[7] A. Ruggiero, J.D. Dattelbaum, A. Pennacchio, L. Iozzino, M. Staiano, M.S. Luchansky, B.S. Der, R. Berisio, S. D'Auria, L. Vitagliano, Crystallization and preliminary X-ray crystallographic analysis of ligand-free and arginine-bound forms of Thermotoga maritima arginine-binding protein, Acta Crystallogr Sect F Struct Biol Cryst Commun 67(Pt 11) (2011) 1462-5.
[8] A. Ausili, A. Pennacchio, M. Staiano, J.D. Dattelbaum, D. Fessas, A. Schiraldi, S. D'Auria, Amino acid transport in thermophiles: characterization of an arginine-binding protein from Thermotoga maritima. 3. Conformational dynamics and stability, J Photochem Photobiol B 118 (2013) 66-73.
[9] A. Ausili, M. Staiano, J. Dattelbaum, A. Varriale, A. Capo, S. D'Auria, Periplasmic Binding Proteins in Thermophiles: Characterization and Potential Application of an Arginine-Binding Protein from Thermotoga maritima: A Brief Thermo-Story, Life (Basel) 3(1) (2013) 149-60.
[10] L.J. Deacon, H. Billones, A.A. Galyean, T. Donaldson, A. Pennacchio, L. Iozzino, S. D'Auria, J.D. Dattelbaum, Tryptophan-scanning mutagenesis of the ligand binding pocket in Thermotoga maritima arginine-binding protein, Biochimie 99 (2014) 208-14.
[11] A. Ruggiero, J.D. Dattelbaum, M. Staiano, R. Berisio, S. D'Auria, L. Vitagliano, A loose domain swapping organization confers a remarkable stability to the dimeric structure of the arginine binding protein from Thermotoga maritima, PLoS One 9(5) (2014) e96560.
[12] G. Gotte, D.V. Laurents, A. Merlino, D. Picone, R. Spadaccini, Structural and functional relationships of natural and artificial dimeric bovine ribonucleases: new scaffolds for potential antitumor drugs, FEBS Lett 587(22) (2013) 3601-8.
[13] N.M. Mascarenhas, S. Gosavi, Understanding protein domain-swapping using structure-based models of protein folding, Prog Biophys Mol Biol 128 (2017) 113-120.
[14] F. Rousseau, J. Schymkowitz, L.S. Itzhaki, Implications of 3D domain swapping for protein folding, misfolding and function, Adv Exp Med Biol 747 (2012) 137-52.
[15] Y. Liu, D. Eisenberg, 3D domain swapping: as domains continue to swap, Protein Sci 11(6) (2002) 1285-99.
[16] L. Mazzarella, L. Vitagliano, A. Zagari, Swapping Structural Determinants of Ribonucleases - an Energetic Analysis of the Hinge Peptide-16-22, Proc Natl Acad Sci U S A 92(9) (1995) 3799-3803.
[17] L. Vitagliano, A. Merlino, A. Zagari, L. Mazzarella, Productive and nonproductive binding to ribonuclease A: X-ray structure of two complexes with uridylyl(2 ',5 ')guanosine, Protein Science 9(6) (2000) 1217-1225.
[18] A. Merlino, L. Vitagliano, M.A. Ceruso, L. Mazzarella, Dynamic properties of the N-terminal swapped dimer of ribonuclease A, Biophys J 86(4) (2004) 2383-2391.
[19] G. Smaldone, M. Vigorita, A. Ruggiero, N. Balasco, J.D. Dattelbaum, S. D'Auria, P. Del Vecchio, G. Graziano, L. Vitagliano, Proline 235 plays a key role in the regulation of the oligomeric states of Thermotoga maritima Arginine Binding Protein, Biochim Biophys Acta 1864(7) (2016) 814-24.

[20] T. Donaldson, L. Iozzino, L.J. Deacon, H. Billones, A. Ausili, S. D'Auria, J.D. Dattelbaum, Engineering a switch-based biosensor for arginine using a Thermotoga maritima periplasmic binding protein, Anal Biochem 525 (2017) 60-66.
[21] M. Staiano, A. Pennacchio, A. Varriale, A. Capo, A. Majoli, C. Capacchione, S. D'Auria, Enzymes as Sensors, Methods Enzymol 589 (2017) 115-131.
[22] G. Smaldone, R. Berisio, N. Balasco, S. D'Auria, L. Vitagliano, A. Ruggiero, Domain swapping dissection in Thermotoga maritima Arginine Binding Protein: how structural flexibility may compensate destabilization, Biochim Biophys Acta In the press (2018).
[23] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol 276 (1997) 307-26.
[24] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J Appl Crystallogr 40(Pt 4) (2007) 658-674.
[25] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr D Biol Crystallogr 53(Pt 3) (1997) 240-55.
[26] R.J. Morris, A. Perrakis, V.S. Lamzin, ARP/wARP and automatic interpretation of protein electron density maps, Methods Enzymol 374 (2003) 229-44.
[27] P.A. Karplus, Experimentally observed conformation-dependent geometry and hidden strain in proteins, Protein science : a publication of the Protein Society 5(7) (1996) 1406-20.
[28] L. Esposito, L. Vitagliano, A. Zagari, L. Mazzarella, Pyramidalization of backbone carbonyl carbon atoms in proteins, Protein science : a publication of the Protein Society 9(10) (2000) 2038-42.
[29] R. Improta, L. Vitagliano, L. Esposito, Peptide bond distortions from planarity: new insights from quantum mechanical calculations and peptide/protein crystal structures, PLoS One 6(9) (2011) e24533.
[30] L. Esposito, N. Balasco, A. De Simone, R. Berisio, L. Vitagliano, Interplay between peptide bond geometrical parameters in nonglobular structural contexts, Biomed Res Int 2013 (2013) 326914.
[31] N. Balasco, L. Esposito, L. Vitagliano, Factors affecting the amplitude of the tau angle in proteins: a revisitation, Acta Crystallogr D Struct Biol 73(Pt 7) (2017) 618-625.
[32] N. Balasco, L. Esposito, A.S. Thind, M.R. Guarracino, L. Vitagliano, Dissection of Factors Affecting the Variability of the Peptide Bond Geometry and Planarity, Biomed Res Int 2017 (2017) 2617629.
[33] P.L. Privalov, Stability of proteins: small globular proteins, Adv Protein Chem 33 (1979) 167-241.
[34] Y. Zhou, C.K. Hall, M. Karplus, The calorimetric criterion for a two-state process revisited, Protein Sci 8(5) (1999) 1064-74.
[35] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J Comput Chem 26(16) (2005) 1701-18.
[36] A. Merlino, L. Vitagliano, M.A. Ceruso, L. Mazzarella, Subtle functional collective motions in pancreatic-like ribonucleases: from ribonuclease A to angiogenin, Proteins 53(1) (2003) 101-10.
[37] A. Amadei, M.A. Ceruso, A. Di Nola, On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations, Proteins 36(4) (1999) 419-24.
[38] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J Mol Graph 14(1) (1996) 33-8, 27-8.
[39] S.E. Zale, A.M. Klibanov, Why does ribonuclease irreversibly inactivate at high temperatures?, Biochemistry 25(19) (1986) 5432-44.
[40] F. Catanzano, A. Gambuti, G. Graziano, G. Barone, Interaction with D-glucose and thermal denaturation of yeast hexokinase B: A DSC study, J Biochem 121(3) (1997) 568-77.
[41] P. Del Vecchio, G. Graziano, G. Barone, L. Mandrich, M. Rossi, G. Manco, Temperature-induced denaturation of Aes acetyl-esterase from Escherichia coli, Thermochim Acta 441(2) (2006) 144-149.
[42] F. Foglia, L. Mandrich, M. Pezzullo, G. Graziano, G. Barone, M. Rossi, G. Manco, P. Del Vecchio, Role of the N-terminal region for the conformational stability of esterase 2 from Alicyclobacillus acidocaldarius, Biophys Chem 127(1-2) (2007) 113-22.
[43] J.F. Brandts, C.Q. Hu, L.N. Lin, M.T. Mos, A simple model for proteins with interacting domains. Applications to scanning calorimetry data, Biochemistry 28(21) (1989) 8588-96.

[44] D.C. Rees, A.D. Robertson, Some thermodynamic implications for the thermostability of proteins, Protein Sci 10(6) (2001) 1187-94.
[45] I.N. Berezovsky, W.W. Chen, P.J. Choi, E.I. Shakhnovich, Entropic stabilization of proteins and its proteomic consequences, PLoS Comput Biol 1(4) (2005) e47.
[46] A. Karshikoff, L. Nilsson, R. Ladenstein, Rigidity versus flexibility: the dilemma of understanding protein thermal stability, Febs J 282(20) (2015) 3899-917.
[47] A. Pica, G. Graziano, Shedding light on the extra thermal stability of thermophilic proteins, Biopolymers 105(12) (2016) 856-63.
[48] C.B. Felder, R.C. Graul, A.Y. Lee, H.P. Merkle, W. Sadee, The Venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors, AAPS PharmSci 1(2) (1999) E2.
[49] N. Verma, A.K. Singh, M. Singh, L-arginine biosensors: A comprehensive review, Biochem Biophys Rep 12 (2017) 228-239.
[50] J.H. Whitfield, W.H. Zhang, M.K. Herde, B.E. Clifton, J. Radziejewski, H. Janovjak, C. Henneberger, C.J. Jackson, Construction of a robust and sensitive arginine biosensor through ancestral protein reconstruction, Protein Sci 24(9) (2015) 1412-22.

Figure legends

Figure 1. Structure and sequence of the wild-type TmArgBP$^{20-246}$. The cartoons of the ligand-free (left – PDB code 4PRS) and arginine-bound (central – PDB code 4PSH) forms of the protein are reported. A single chain of the dimer is reported in foreground whereas the other is in background as transparent cartoon. For the foreground subunit the following color code has been used: D1 in blue, D2 in green and the swapping C-terminal helix in red. The protein residues directly involved in arginine binding are also shown. The arginine-binding residues Gln142, Thr146, and Asp183 belong to the domain 2. All the others are of D1. The sequence of the wild-type protein deprived of the trans membrane N-terminal fragment (residues 1-19) is reported at the bottom of the figure using the same color code adopted for the structures.

Figure 2. SEC-LS analysis of D1 (A) and D2 (B) domains. The experiments were performed using a Superdex S200 10/30 column pre-equilibrated with a buffer containing Tris-HCl 50 mM (pH 8) and NaCl 150 mM.

Figure 3. Far-UV CD spectra of D1 (A) and D2 (B) domains. The spectra were recorded using a protein concentration of 20 μM.

Figure 4. Thermal denaturation curves of D1 (A) and D2 (B) domains. The protein was dissolved in PBS buffer (pH 7.4) containing 4M GuHCl.

Figure 5. Titration of arginine-bound D1 with D2 (A) or buffer (B) followed by ITC. The subtraction of the two experiments is reported in panel C.

Figure 6. Titration of arginine with D1 (A) and D2 (B) domains followed by ITC. Top and bottom panel reports raw and integrated data, respectively.

Figure 7. Crystal structure of D1 (blue) and D2 (green) superimposed to the structure of the ligand-free TmArgBP$^{20-233}$ (grey).

Figure 8. Superimposition of the structures of ligand-free (cyan) and arginine-bound (pink) D1 domain. The ligand arginine of the bound form and the residues of the binding pockets of the two domains are shown.

Figure 9. Crystal structure of the D2 domain. The residues involved in arginine recognition in the wild-type TmArgBP are shown.

Figure 10. DSC traces of the arginine-bound (black line) and ligand-free (red line) forms of TmArgBP$^{20-233}$ at pH 7.4 (A). DSC trace of the arginine-bound form of TmArgBP$^{20-233}$ at pH 7.4, first heating stopped at 105 °C (black line) and second heating up to 120 °C (red line) (B).

Figure 11. DSC traces of the ligand-free (red line) and arginine-bound (black line) forms of D1 at pH 7.4 (A). DSC trace of the arginine-bound form of D1 domain at pH 7.4, first heating stopped at 95 °C (black line) and second heating up to 110 °C (red line) (B).

Figure 12. DSC trace of D2 domain at pH 7.4 (A). DSC trace of D2 domain at pH 7.4, first heating stopped at 103 °C (black line) and second heating up to 115 °C (red line) (B).

Figure 13. Comparison among the DSC traces of the ligand-free form of TmArgBP$^{20-233}$ (black line), of the arginine-bound form of D1domain (red line) and of D2 domain (blue line) at pH 7.4.

Figure 14. RMSD values, computed on the C^{α} atoms, of trajectory structures compared to the starting crystallographic structures of TmArgBP$^{20-233}$(A), D1 (B) and D2 (C). For TmArgBP$^{20-233}$, the RMSD values of the entire protein (grey) and of the individual domains D1 (blue) and D2 (green) are reported. The time evolution of distances between the centers of mass of D1 and D2 in TmArgBP$^{20-233}$ is reported in panel D.

Figure 15. RMSF values, computed on the C^{α} atoms, in the simulation of D1 (cyan) and D2 (light green). The RMSF values of individual domains D1 (blue) and D2 (green) in TmArgBP$^{20-233}$ context are reported.

Figure 16. Overall scheme of TmArgBP dissection from the wild-type dimer to the individual domains.

Table 1. Data collection statistics. Values in parentheses are for the highest resolution shells (1.78-1.75 Å, 1.81-1.75 Å. and 1.77-1.73 for ligand-free D1, arginine-bound D1, and D2, respectively).

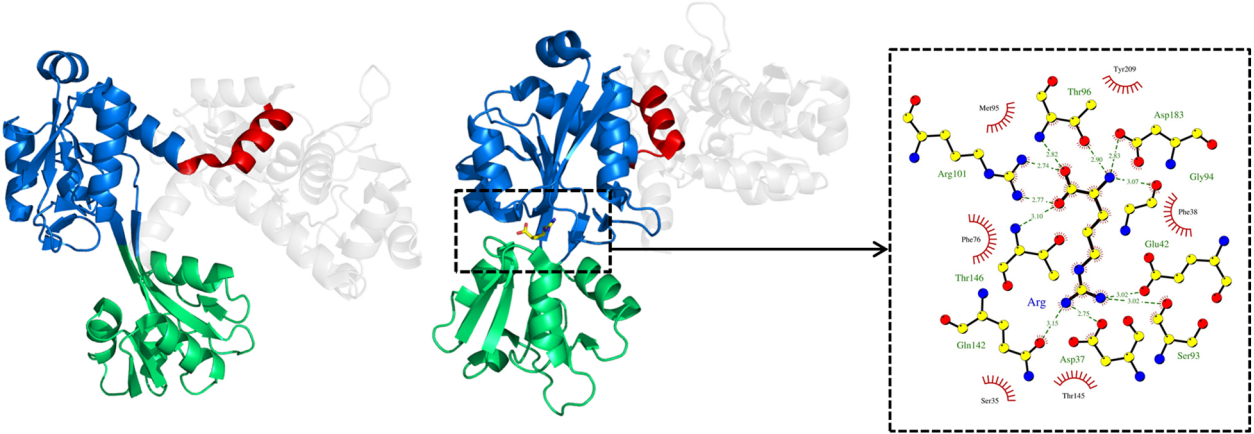
Protein	D1 (ligand-free)	D1 (arginine-bound)	D2
X-ray device	Rigaku FR007HF with CCD detector	Rigaku FR007HF with CCD detector	Rigaku FR007HF with CCD detector
Space group	P1	P1	P21
Unit cell parameters			
a (Å)	35.45	35.53	25.62
b (Å)	46.04	45.93	49.73
c (Å)	52.50	52.51	27.78
α (°)	116.02	115.91	90.0
β (°)	109.85	109.53	100.19
γ (°)	90.01	90.02	90.00
Resolution range (Å)	50.00-1.75	50.00-1.75	50.00-1.73
Wavelength (Å)	1.54	1.54	1.54
Average redundancy	2.7 (1.9)	2.8 (1.8)	1.3 (1.0)
Unique reflections	24812	25022	6898
Completeness (%)	88.8 (67.3)	89.4 (71.3)	95.1 (75.6)
R merge (%)	4.4 (8.5)	6.2 (22.8)	4.0 (3.0)
Average I/σ(I)	37.2 (9.2)	22.1 (3.3)	30.2 (16.8)

Table 2. Refinement statistics

Protein	D1 (ligand-free)	D1 (arginine-bound)	D2
Asymmetric unit	2 molecules	2 molecules	1 molecule
Resolution (Å)	14.40-1.75	14.62-1.75	14.18-1.73
R, Rfree (%)	16.4, 18.8	16.2, 20.2	12.9, 16.8
No. of residues	252	252	92
Mean B value ($Å^2$)	18.3	17.3	16.7
R.m.s. deviations			
Bond lengths (Å)	0.021	0.019	0.014
Bond angles (°)	1.95	2.04	1.67

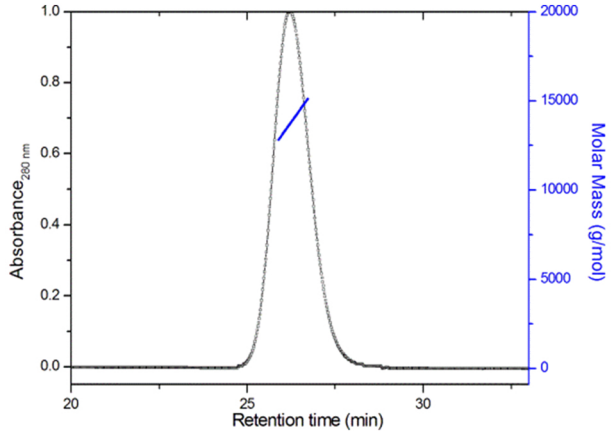
Table 3. Experimental values of the denaturation temperature, denaturation enthalpy change, cooperative unit, CU = $\Delta_d H(T_d)/\Delta_d H(T_d)^{vH}$, and van't Hoff enthalpy change for the investigated proteins at pH 7.4, PBS buffer.

	T_d (°C)	$\Delta_d H(T_d)$ (kJ mol^{-1})	CU	$\Delta_d H^{vH}$ (kJ mol^{-1})
TmArgBP$^{20-233}$ (ligand-free)	108.5	1110	0.97	1150
TmArgBP$^{20-233}$ (arginine-bound)	105.3	1000	1.0	995
D1 (ligand-free)	95.5	450	0.98	460
D1 (arginine-bound)	95.6	450	0.97	465
D2	102.6	470	0.96	490

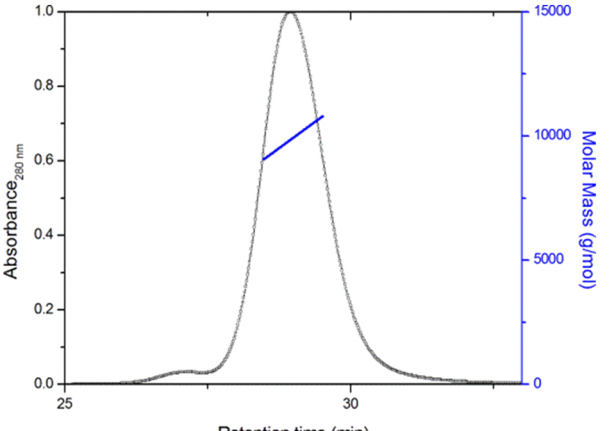


20_AIDEIKSRGYLLVGLSADFPPFEFVDENGNIVGFDVDLAKEIARRLGVELKIVDMTFDGLIPSLLTKKIDVIISG
MTITEERKKVVAFSDPYFDAGQVIVVRKDSDFRPKTYEDLVGKTVAVQIGTTGDIEVSKYDGIKVVRFDKFTDAFLEL
KRGRADAVVLDSATARAFVAKNPDLVISSGVLSSEQYGIAVRKEDTDLLEFINSVLRELKKSPYDVLIEKWFSE_246

Figure 1

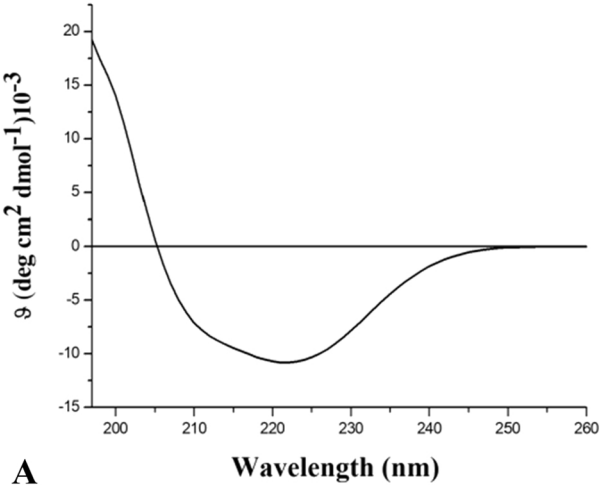

1.0
0.8
0.6
0.4
0.2
0.0
Absorbance280 nm
20000
15000
10000
5000
0
Molar Mass (g/mol)
20
25
30
Retention time (min)

A

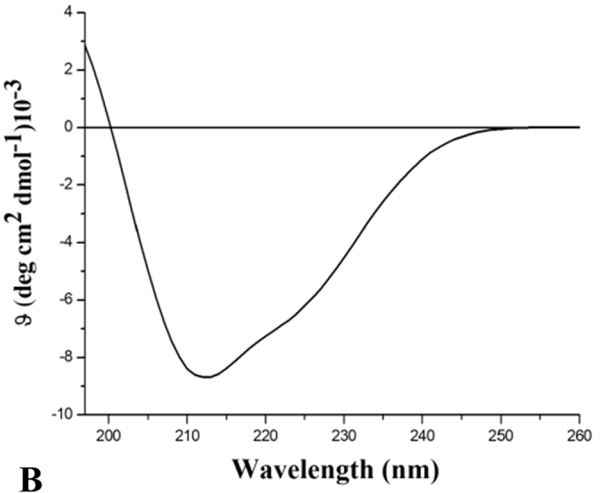

1.0
0.8
0.6
0.4
0.2
0.0
Absorbance280 nm
15000
10000
5000
0
Molar Mass (g/mol)
25
30
Retention time (min)

B

Figure 2

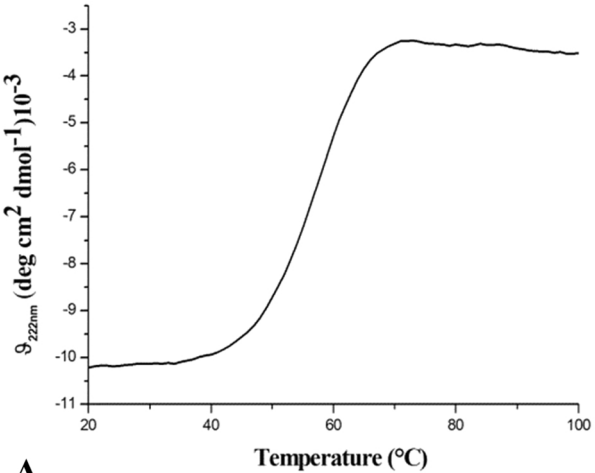

ϑ (deg cm2 dmol-1)10-3
20
15
10
5
0
-5
-10
-15
200
210
220
230
240
250
260
Wavelength (nm)

A


ϑ (deg cm2 dmol-1)10-3
4
2
0
-2
-4
-6
-8
-10
200
210
220
230
240
250
260
Wavelength (nm)

B

Figure 3


ϑ_{222nm} (deg cm^2 dmol^{-1})10^{-3}
-3
-4
-5
-6
-7
-8
-9
-10
-11
20
40
60
80
100
Temperature (°C)
A

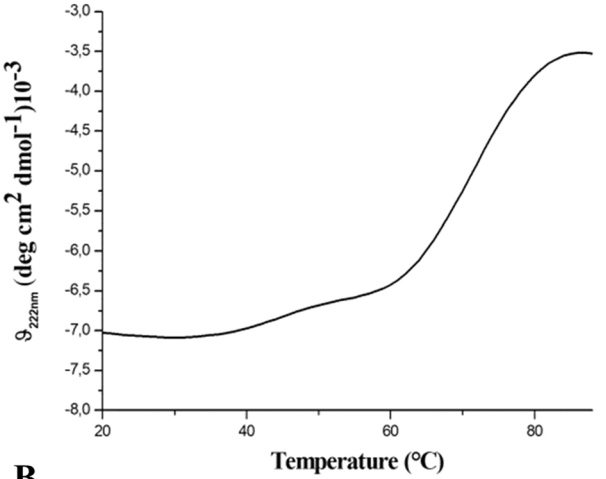
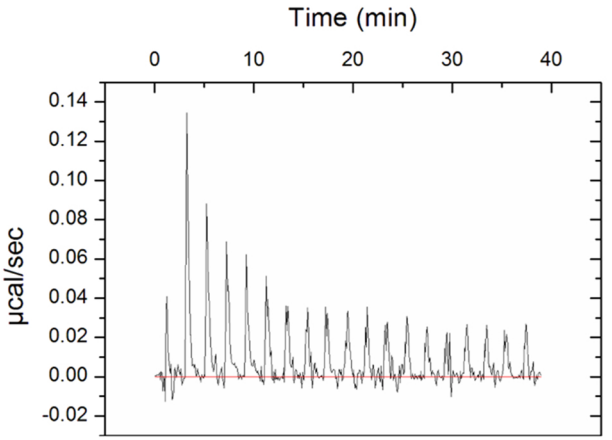
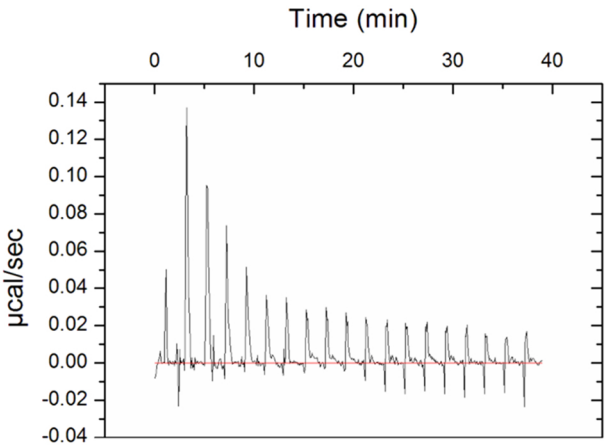

ϑ_{222nm} (deg cm^2 dmol^{-1})10^{-3}
-3,0
-3,5
-4,0
-4,5
-5,0
-5,5
-6,0
-6,5
-7,0
-7,5
-8,0
20
40
60
80
Temperature (°C)
B

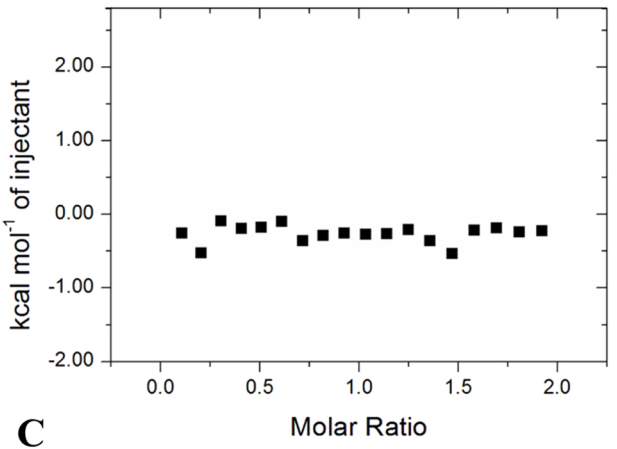
Figure 4


Time (min)
0
10
20
30
40
µcal/sec
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02

A


Time (min)
0
10
20
30
40
µcal/sec
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
-0.04

B


kcal mol-1 of injectant
2.00
1.00
0.00
-1.00
-2.00
0.0
0.5
1.0
1.5
2.0
Molar Ratio

C

Figure 5

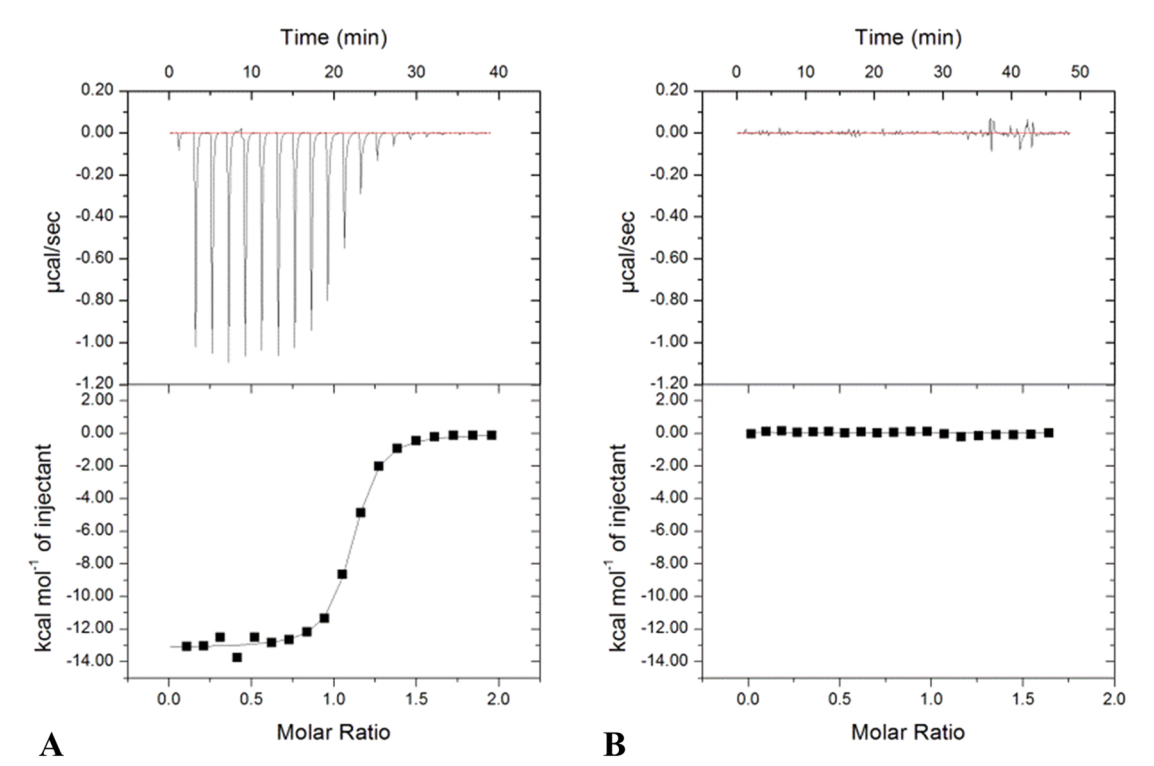

Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
A
Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
B

Figure 6

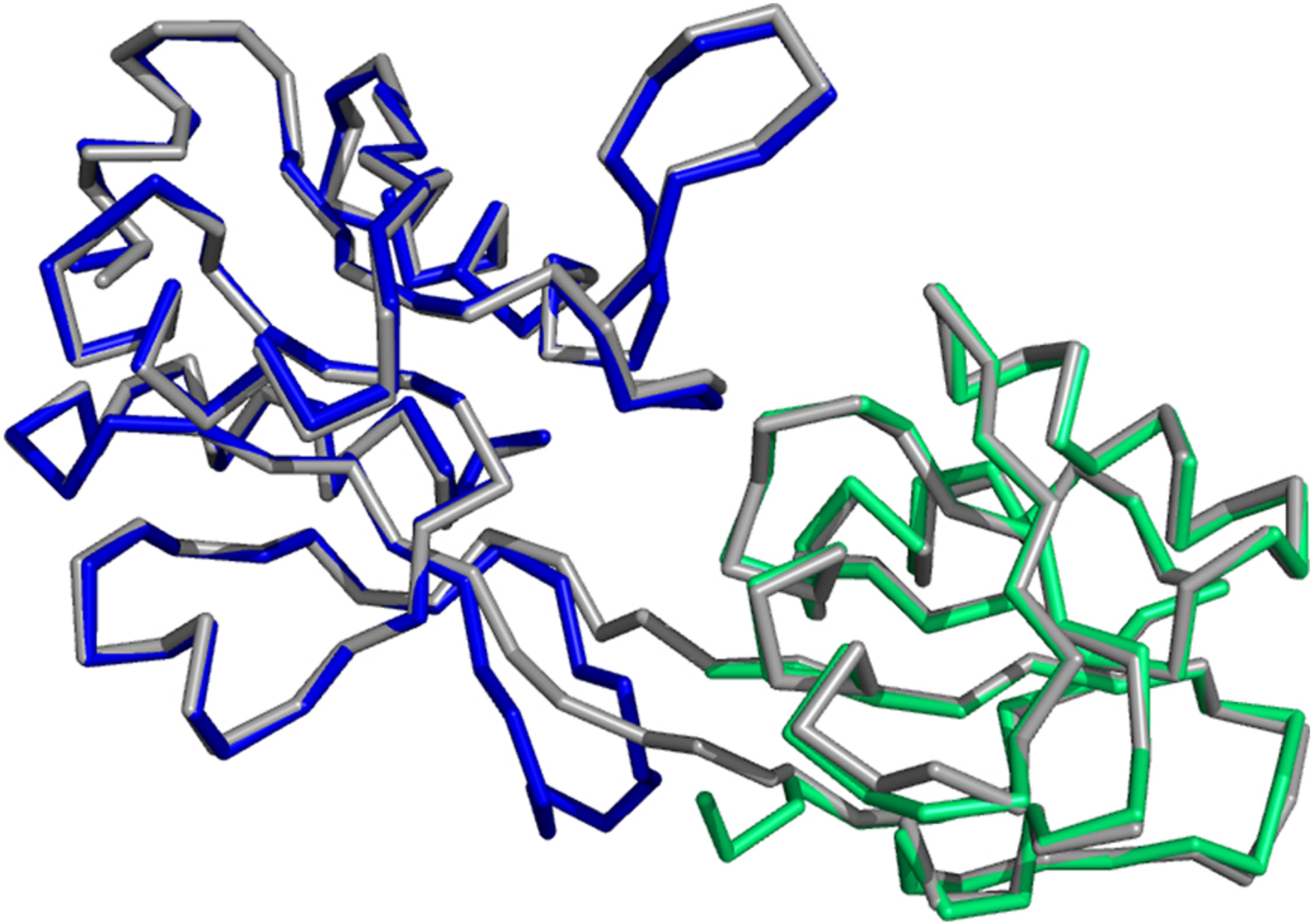


Figure 7

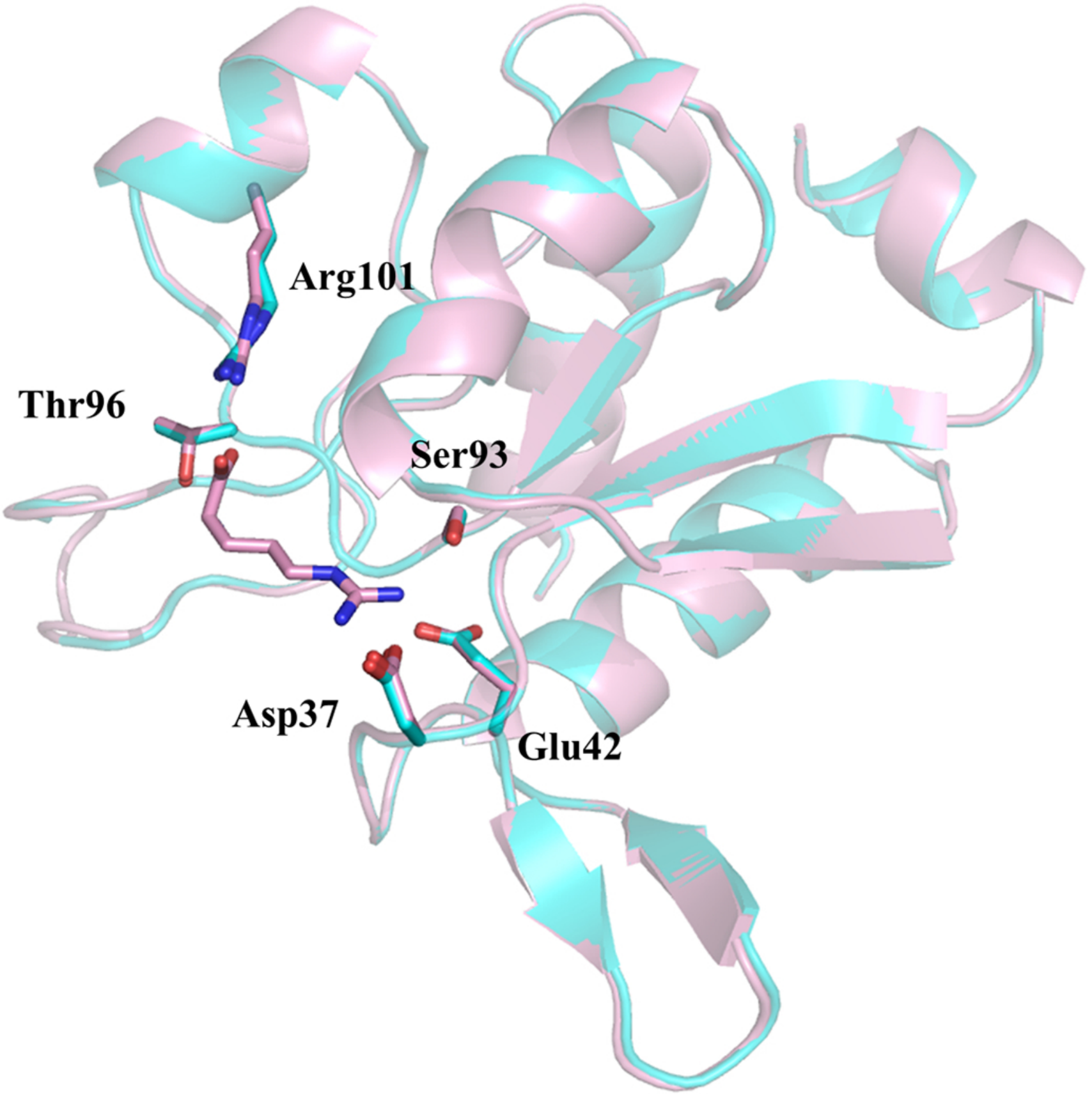


Figure 8

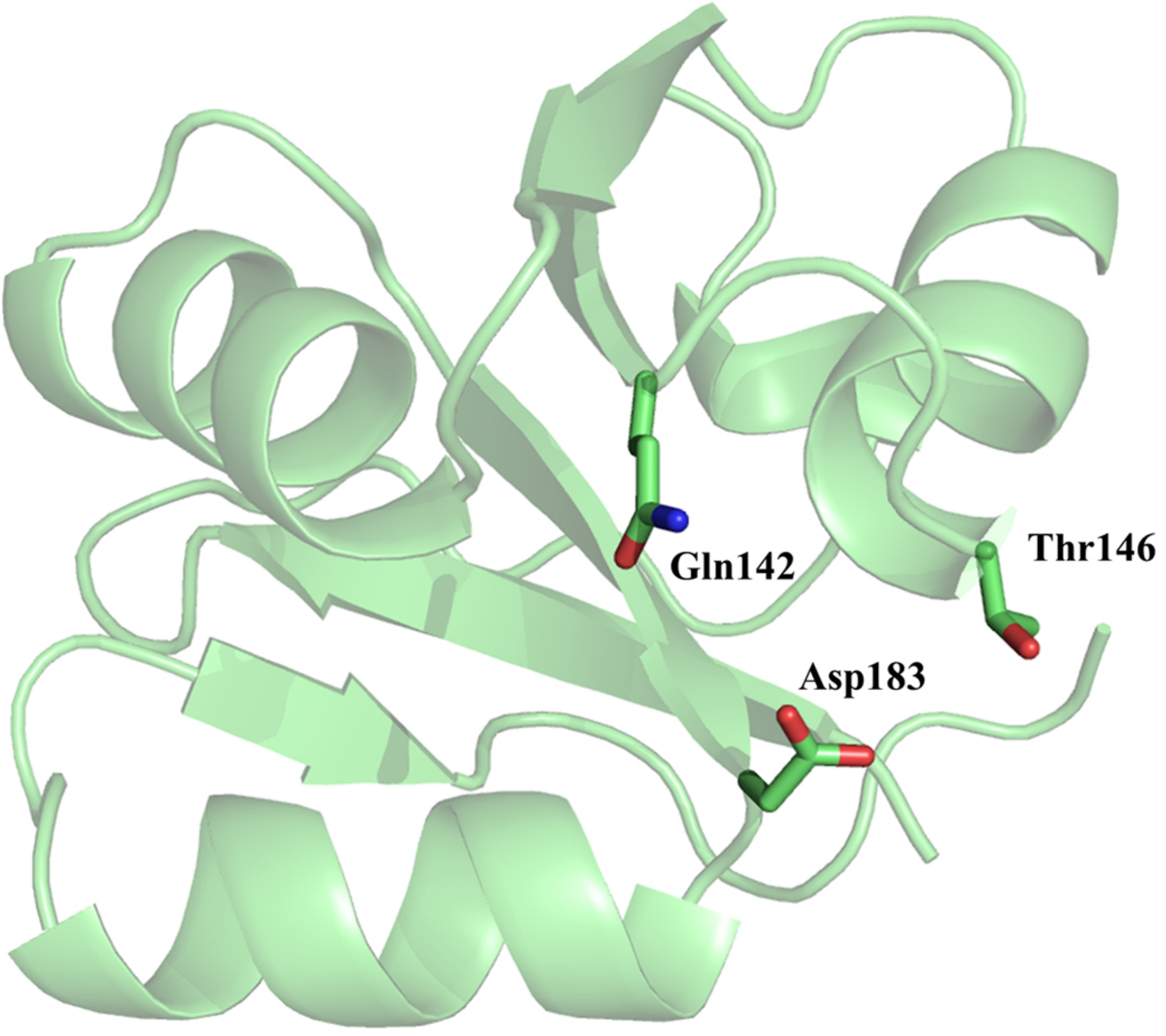


Figure 9

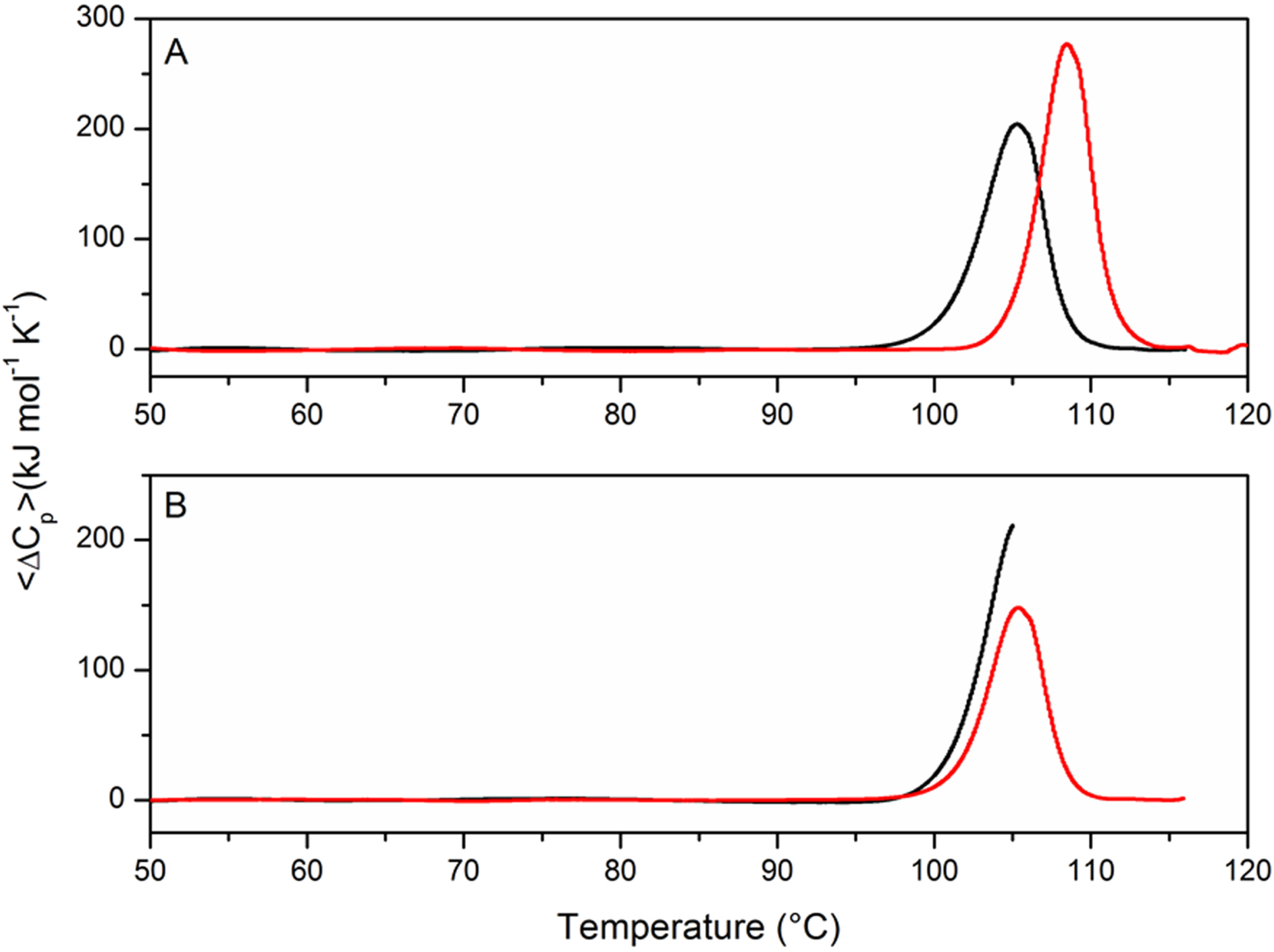

A
B
300
200
100
0
50
60
70
80
90
100
110
120
<ΔCp>(kJ mol-1 K-1)
Temperature (°C)

Figure 10

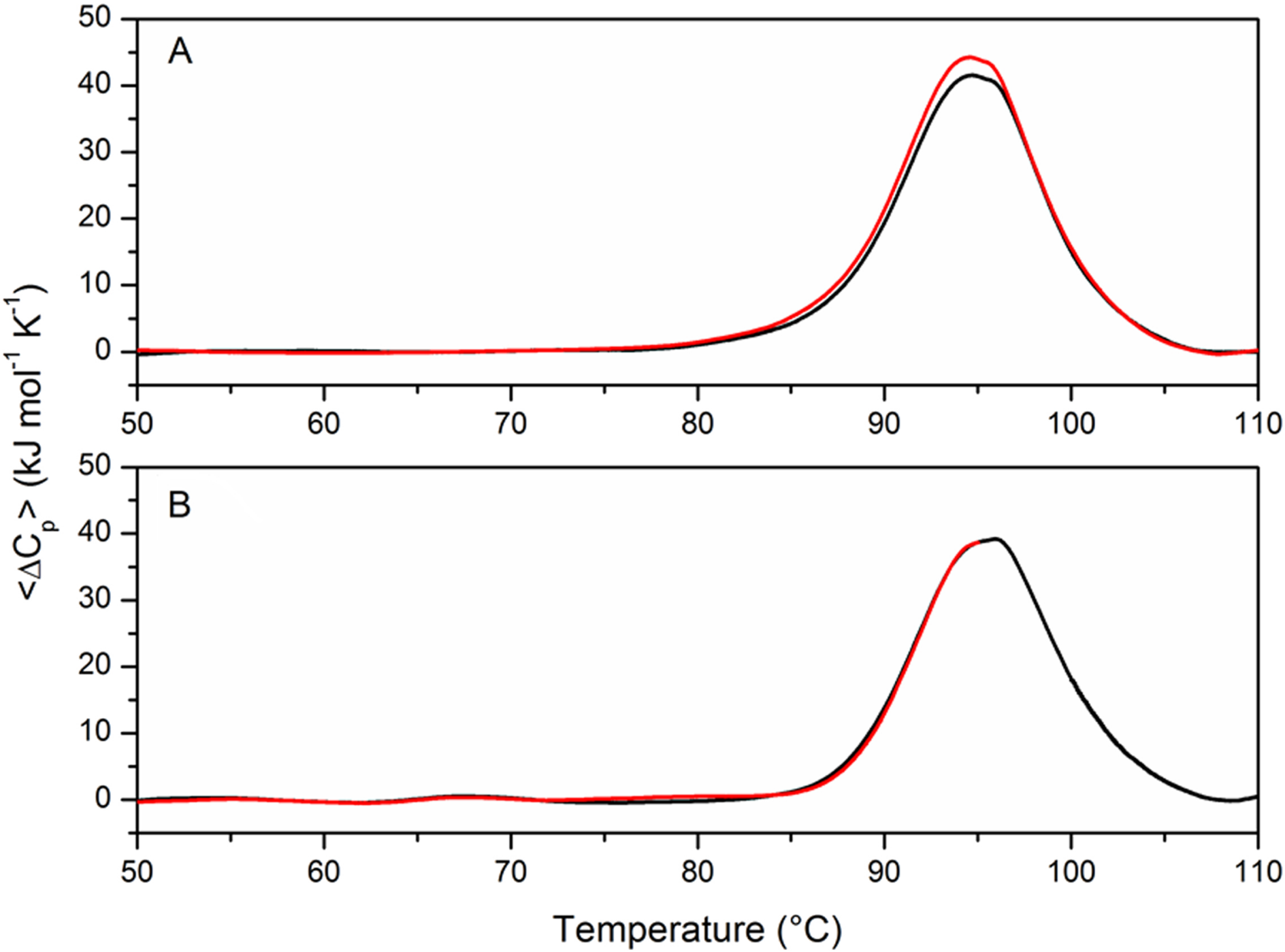

A
B
50
40
30
20
10
0
50
60
70
80
90
100
110
<ΔC_p> (kJ mol^-1 K^-1)
Temperature (°C)

Figure 11

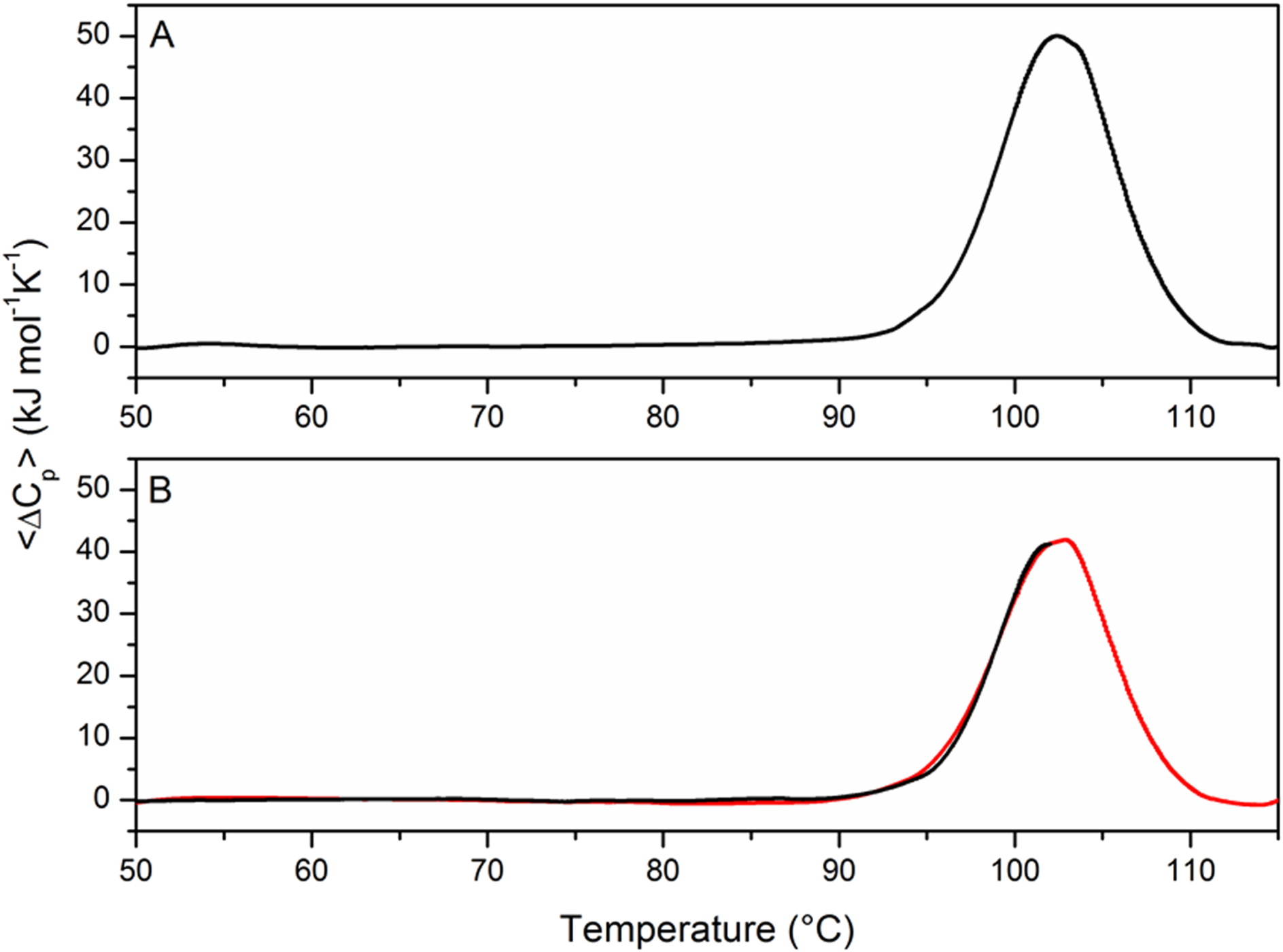

A
B
50
40
30
20
10
0
50
60
70
80
90
100
110
<ΔC_p> (kJ mol^-1 K^-1)
Temperature (°C)

Figure 12

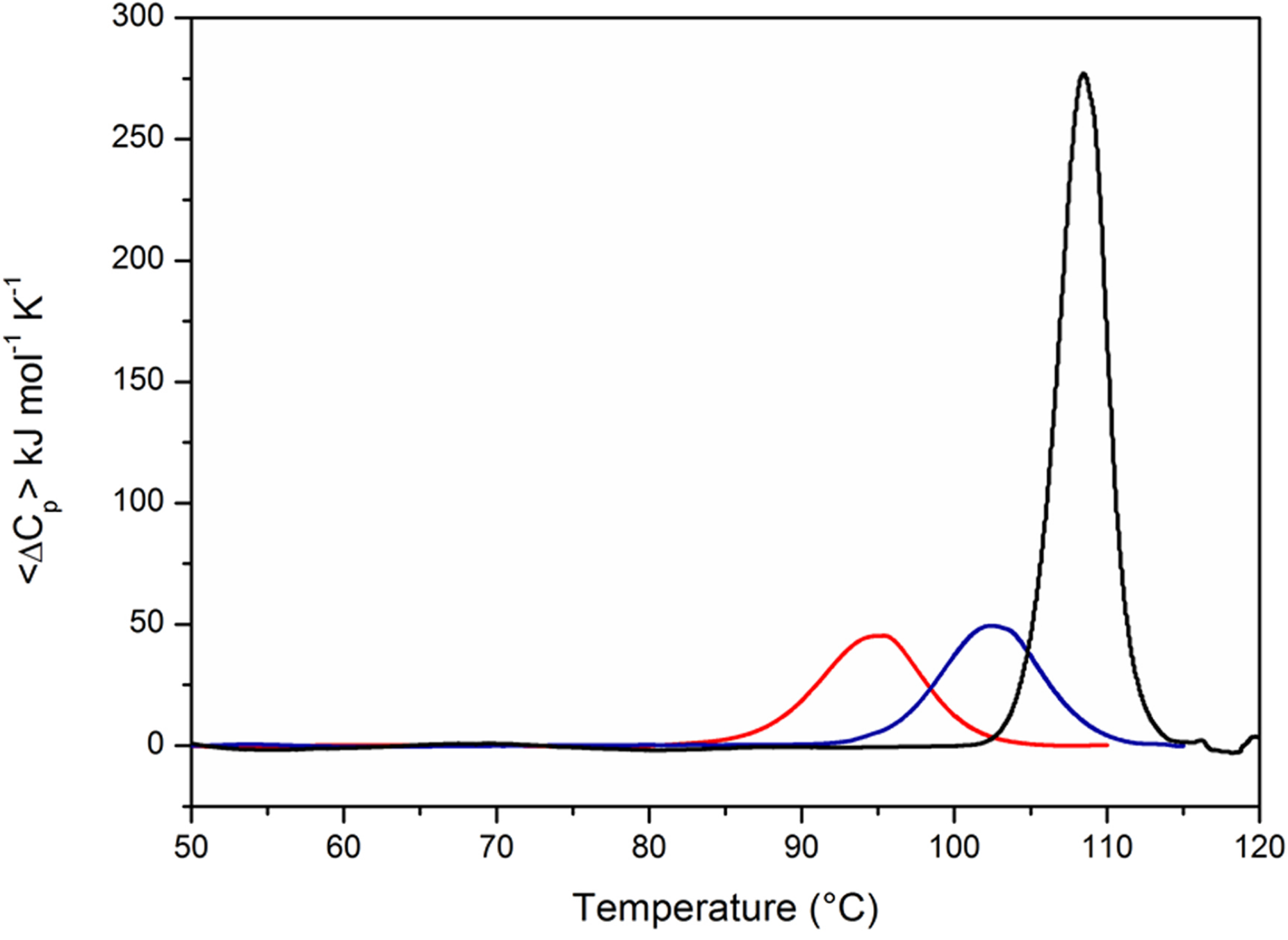
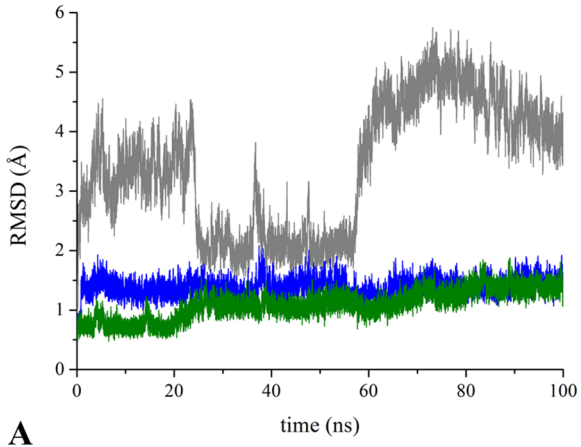
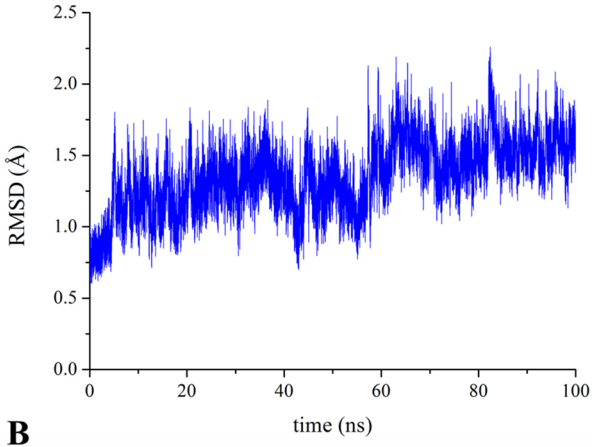
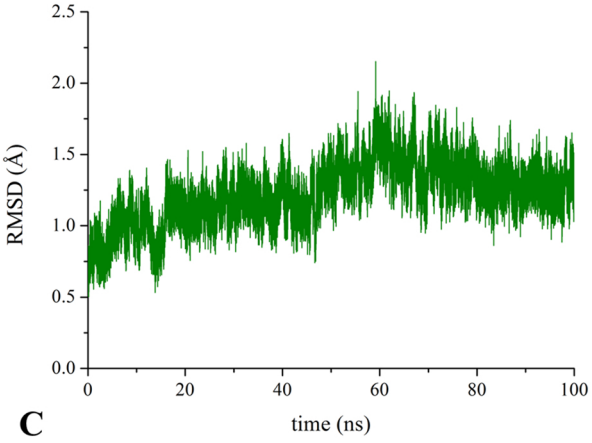

300
250
200
150
100
50
0
<ΔC_p> kJ mol^-1 K^-1
50
60
70
80
90
100
110
120
Temperature (°C)

Figure 13


RMSD (Å)
time (ns)
A


RMSD (Å)
time (ns)
B


RMSD (Å)
time (ns)
C

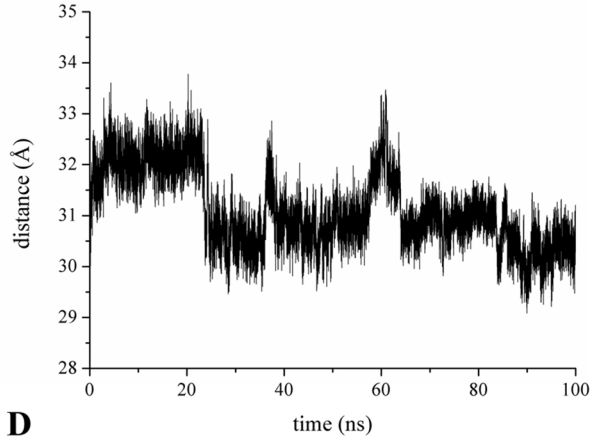

distance (Å)
time (ns)
D

Figure 14

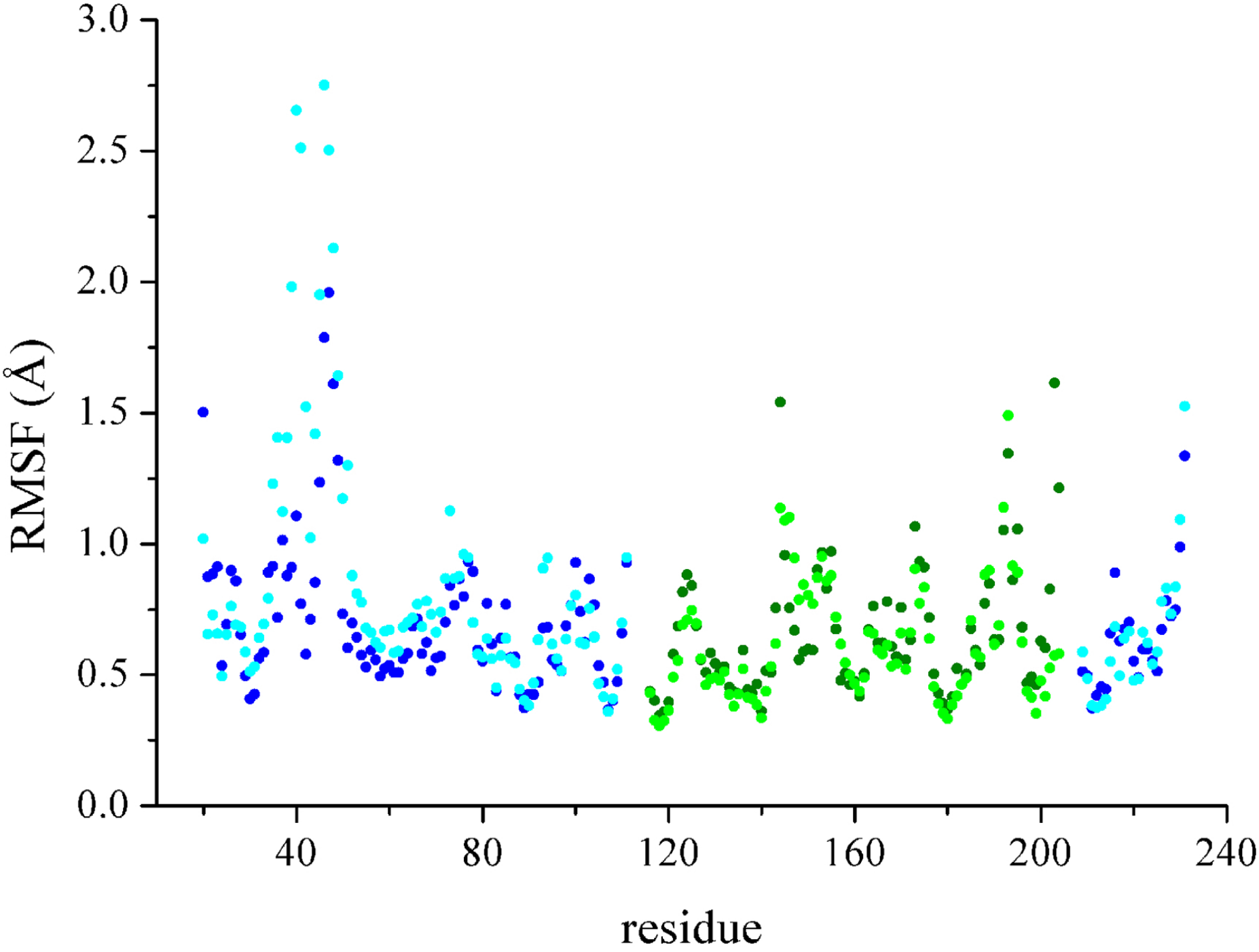


Figure 15

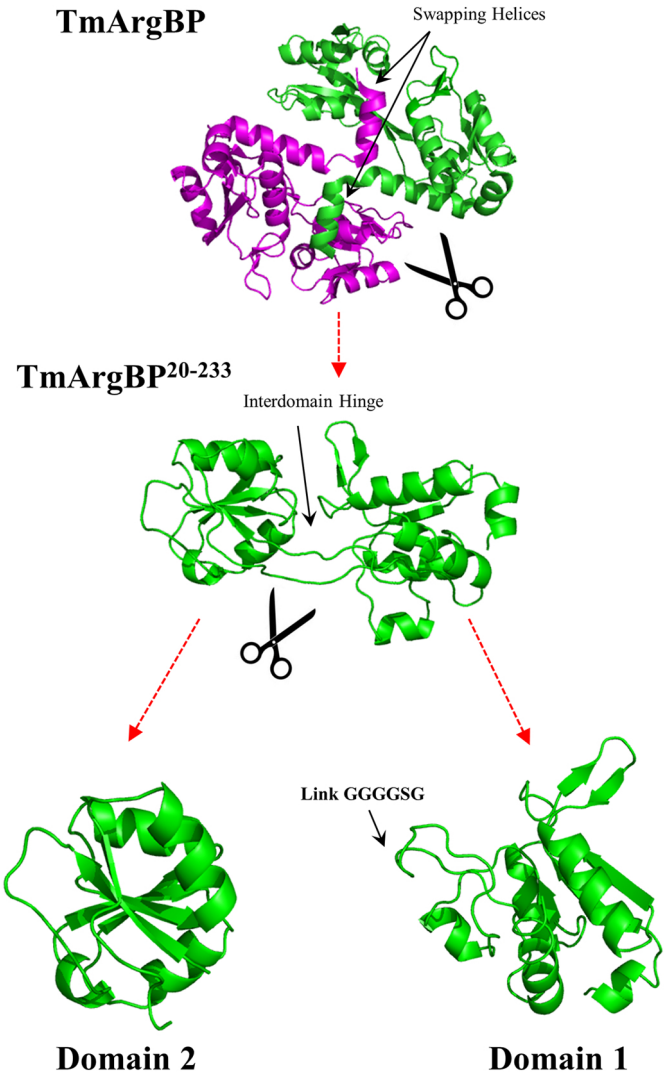

TmArgBP
Swapping Helices
TmArgBP20-233
Interdomain Hinge
Link GGGGSG
Domain 2
Domain 1

Figure 16