The FEBS Journal

FEBS PRESS science publishing by scientists

Unusual dimerization of a *Bc*Csp mutant leads to reduced conformational dynamics

Alonso I. Carvajal[1], Gabriel Vallejos[1], Elizabeth A. Komives[2], Víctor Castro-Fernández[1], Diego A. Leonardo[3], Richard C. Garratt[3], César A. Ramírez-Sarmiento[1,4] and Jorge Babul[1]

1 Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Santiago, Chile
2 Department of Chemistry & Biochemistry, University of California San Diego, La Jolla, CA, USA
3 Instituto de Física de São Carlos, Universidade de São Paulo, Brazil
4 Institute for Biological and Medical Engineering, Schools of Engineering, Medicine and Biological Sciences, Pontificia Universidad Católica de Chile, Santiago, Chile

Keywords
cold shock proteins; conformational dynamics; protein folding; protein-protein interactions; RNA chaperones

Correspondence
C. A. Ramírez-Sarmiento, Institute for Biological and Medical Engineering, Schools of Engineering, Medicine and Biological Sciences, Pontificia Universidad Católica de Chile. Avda. Vicuña Mackenna 4860, Macul, Santiago, Chile
Tel: +56 2 2354 1110
E-mail: cesar.ramirez@uc.cl
Jorge Babul, Laboratorio de Bioquímica y Biología Molecular, Departamento de Biología, Facultad de Ciencias, Universidad de Chile. Las Palmeras 3425, Ñuñoa, Santiago 7800003, Chile
Tel: +56 2 2978 7231
E-mail: jbabul@uchile.cl

(Received 3 March 2017, revised 9 April 2017, accepted 26 April 2017)

doi:10.1111/febs.14093

Cold shock proteins (Csp) constitute a family of ubiquitous small proteins that act as RNA-chaperones to avoid cold-induced termination of translation. All members contain two subdomains composed of 2 and 3 β-strands, respectively, which are connected by a hinge loop and fold into a β-barrel. *Bacillus caldolyticus* Csp (*Bc*Csp) is one of the most studied members of the family in terms of its folding, function, and structure. This protein has been described as a monomer in solution, although a recent crystal structure showed dimerization via domain swapping (DS). In contrast, other cold shock proteins of the same fold are known to dimerize in a non-swapped arrangement. Hypothesizing that reducing the size of the hinge loop may promote swapping as in several other DS proteins with different folds we deleted two residues from these region (*Bc*CspΔ36-37), leading to a protein in monomer–dimer equilibrium with similar folding stability to that of the wild-type. Strikingly, the crystal structure of *Bc*CspΔ36-37 revealed a nonswapped dimer with its interface located at the nucleic acid-binding surface, showing that the deletion led to structural consequences far from the perturbation site. Concomitantly, circular dichroism experiments on *Bc*CspΔ36-37 demonstrated that binding of the oligonucleotide hexathymidine disrupts the dimer. Additionally, HDXMS shows a protective effect on the protein structure upon dimerization, where the resulting interactions between ligand-binding surfaces in the dimer reduced the extent of exchange throughout the whole protein. Our work provides evidence of the complex interplay between conformational dynamics, deletions, and oligomerization within the Csp protein family.

Databases
Structural data are available in the Protein Data Bank under accession number 5JX4.

Introduction

A cold shock response is an adaptive mechanism triggered in bacterial cells when their growth is challenged by a temperature decrease, enabling them to survive in the new environmental conditions [1]. During this

Abbreviations
*Bc*Csp, *Bacillus caldolyticus* Csp; Csp, cold shock protein; HDX, hydrogen/deuterium exchange; HDXMS, hydrogen/deuterium exchange mass spectrometry; ln(Pf), natural logarithm of protection factor; Pf, protection factor; Rs, Stokes radius; SEC, size-exclusion chromatography.

response, cell growth is arrested and the expression of most cellular genes is downregulated [2]. A few proteins, the majority of which are nucleic acid-binding proteins involved in RNA metabolism and protein transcription and translation, are transiently induced during cold shock [3,4]. The major cold shock proteins (Csp), a group of small globular proteins, show the highest relative expression during the cold shock response [5].

The Csps constitute a highly conserved family present in more than 400 different bacteria, including hyperthermophilic, thermophilic, mesophilic, and psychrophilic species [6–8], even containing multiple copies of their encoding genes [5]. Their conserved structure consists of 65–70 amino acid residues folded into a closed five-stranded antiparallel β-barrel (Fig. 1) [9] known as the OB-fold domain [10]. This is frequently used by several nonhomologous families for nucleic acid recognition through a conserved binding surface [11], making it an interesting model for the study of protein evolution [12]. In fact, several eukaryotic RNA-binding proteins also possess a Csp-homologous domain, known as a cold shock domain [13], indicating that the OB-fold is one of the most evolutionary conserved nucleic acid-binding structures within prokaryotes and eukaryotes since the origin of single-cell life [14].

The biological function of Csps is supported by two conserved ribonucleoprotein class motifs known as RNP1 and RNP2, which are composed of basic and aromatic residues located in strands β1 and β2, allowing Csps to interact with single-stranded DNA and RNA [15] by binding to stretches of six to seven nucleotides with micromolar to nanomolar affinity and high specificity [16–18]. The high affinity for single-stranded nucleic acids allows Csps to prevent the formation of mispaired RNA duplexes, thus acting as RNA chaperones [19] and preventing transcription pausing and termination during the cold-shock response [20].

The small size and high conservation of Csps make them a unique system to study intriguing evolutionary aspects of protein folding, structure, sequence, and function. The folding of several Csps has been assessed in detail and has been shown to be extremely rapid, occurring on the millisecond timescale, and following a simple two-state N ↔ U model without any folding intermediates [6,21]. Also, several three-dimensional structures of Csps have been solved to date, including *Escherichia coli* CspA (*Ec*CspA) [9,22,23], *Bacillus subtilis* CspB (*Bs*CspB) [18,24–27], *Thermotoga maritima* [28,29], *Listeria monocytogenes* CspA [8], *Salmonella typhimurium* CspE (*St*CspE) [30], *Neisseria meningitidis* (*Nm*Csp) [31] and *Bacillus caldolyticus* (*Bc*Csp) [7,32]. Despite the high sequence and structural conservation of these proteins, their thermal stabilities are consistent with the optimum growth temperatures of their source organisms [6], a variation that is largely conferred by specific clusters of charged residues on the protein surface [33].

Most Csps have been described to exist only as monomers in solution, whereas several proteins presenting the SH3-fold, which is similar to the OB-fold but with a different strand topology [34], are able to form homo and heterodimers through a myriad of dimer orientations [35]. Nevertheless, two Csps have been described to form several types of dimers: *Bs*CspB exists as a dimer formed through an interaction interface on strand β4 in the crystal structure [25] and can form a dimer in solution in the absence of phosphate ions [36]; *St*CspE forms a dimer both in solution and in the crystal structure through its nucleic acid-binding surface [30]; finally, the crystal structures of *Bc*Csp and *Nm*Csp revealed a domain-swapped dimer, in which strands β1–3 are able to associate with strands β4–5 from the adjacent subunit due to changes in the torsion angles of residues E36 and G37 [7,31]. *Ec*CspA forms amyloid fibrils under acidic conditions [37], a protein aggregation process where, in some cases, domain swapping can play an important role such as in the case of ribonuclease A [38].

Intrigued by the increasing evidence of structural polymorphism in Csps and by the fact that the domain-swapped dimer has not been observed in solution, we performed biophysical characterization of *Bc*Csp mutants obtained by single and double deletions of residues E36 and G37, a mutational strategy previously used to favor the formation of domain-swapped oligomers in other proteins [39,40]. These studies, including structural characterization by X-ray crystallography and hydrogen deuterium exchange mass spectrometry (HDXMS) as well as characterization of the folding of the double deletion mutant, revealed that the double deletion led to formation of a poorly stable nonswapped dimer that is maintained through interactions involving the nucleic acid-binding surface. Interestingly, HDXMS also revealed that dimer formation leads to a decrease in flexibility throughout the whole protein. These observed global and local structural changes suggest that flexible loop regions may be gatekeepers of monomer-to-dimer transitions in small single-domain proteins such as the Csp family.

Results and Discussion

The double deletion Δ36-37 forms a soluble dimer without significantly altering the stability and function of *Bc*Csp

*Bc*Csp has been the target of extensive biophysical characterization, consistently being reported as a two-

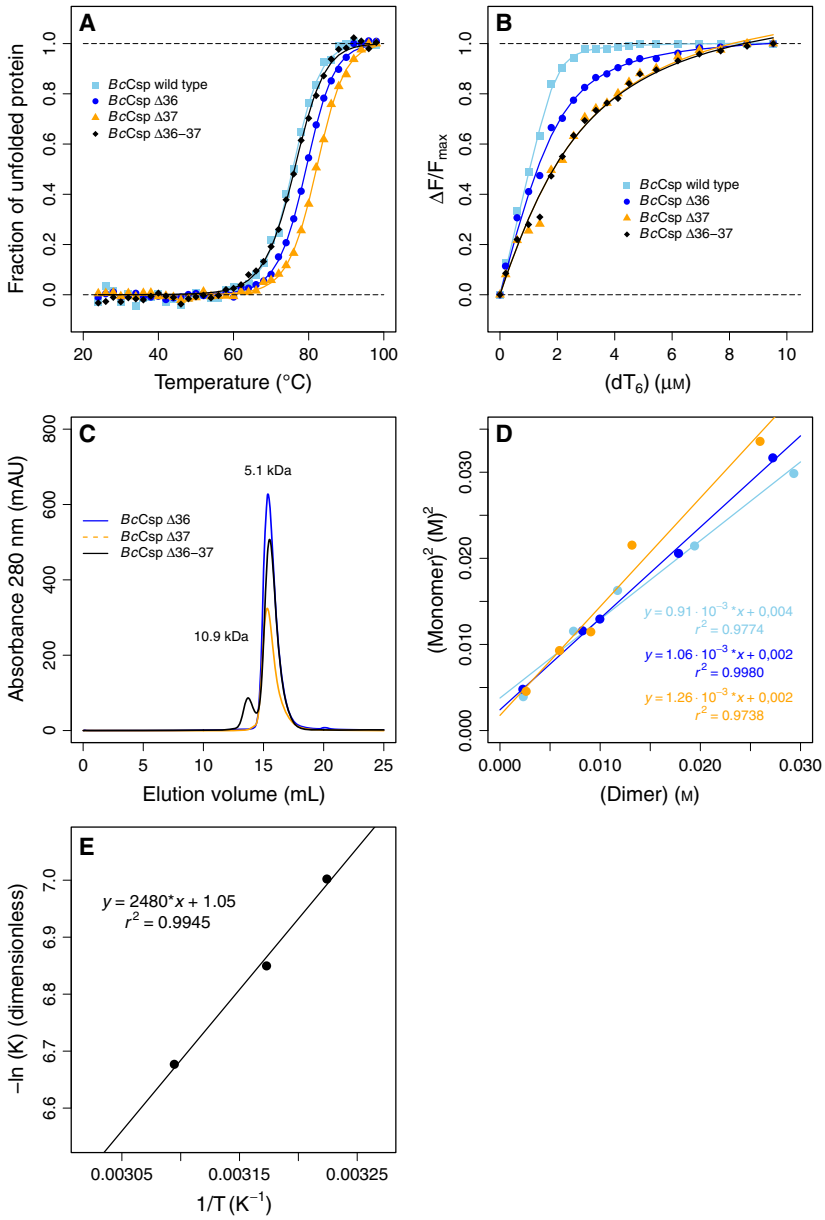

Fig. 1. Biophysical characterization of hinge loop deletion mutants of *Bc*Csp (A) Thermal unfolding of wild-type *Bc*Csp (cyan), *Bc*CspΔ36 (blue), *Bc*CspΔ37 (orange), and *Bc*CspΔ36-37 (black) at 28.2 μM. The normalized fractions of unfolded protein, based on mean residue molar ellipticity obtained at 222 nm, are plotted as a function of the temperature. Data were fitted to a two-state folding mechanism (Eqns 4, 5). (B) Ligand binding of wild-type *Bc*Csp (cyan), *Bc*CspΔ36 (blue), *Bc*CspΔ37 (orange), and *Bc*CspΔ36-37 (black) at 3 μM. Data are shown as normalized fluorescence quenching as a function of the concentration of dT_6. Solid lines represent fitting to the ligand-binding model given in Eqn (8). (C) Oligomeric state of *Bc*CspΔ36 (blue), *Bc*CspΔ37 (orange) and *Bc*CspΔ36-37 (black). The protein concentration was 0.5 mM for *Bc*CspΔ36 and *Bc*CspΔ37, and 1 mM for *Bc*CspΔ36-37. SEC elution profiles were obtained by measuring the absorbance at 280 nm. The estimated molecular weights from the column calibration are indicated on top of each peak. (D) Determination of equilibrium dissociation constants (K_D) of the dimer of *Bc*CspΔ36-37 at three temperatures. Protein concentrations of dimeric and monomeric species of *Bc*CspΔ36-37 were incubated until reaching equilibrium at 37 °C (cyan), 42 °C (blue), and 50 °C (orange). The solid lines represent linear fits of each dataset using Eqns (1, 2). (E) van't Hoff's analysis of the monomer–dimer equilibrium of *Bc*CspΔ36-37. The K_D were extracted from the slopes of the linear fits in D. The solid line represents the linear fit of these data (Eqn 3).

state folding monomer for both wild-type protein [32] and a large number of mutants [33,41,42]. The only exception being a crystallographic domain-swapped dimer of the wild-type protein in the presence of hexathymidine that has not been detected in solution [7]. In other models of domain swapping, limiting the

flexibility of hinge loops through proline substitutions and single-residue deletions has guided folding toward the formation of this type of intertwined oligomer [39,40], sometimes accompanied by changes in protein stability [43].

We assessed the role of flexibility of the hinge region of *Bc*Csp on its folding, function, and oligomerization by thermal unfolding, nucleic acid binding, and SEC experiments of the engineered single and double deletion mutants *Bc*Csp Δ36, Δ37, and Δ36-37. The results from these analyses are presented in Fig. 1. First, we observe that all proteins follow a two-state unfolding mechanism (Fig. 1A). Both single deletion mutants *Bc*CspΔ36 and *Bc*CspΔ37, whose ΔG_U at 70 °C are 1.5 ± 0.14 and 2.0 ± 0.14 kcal·mol^{-1}, respectively, are slightly more stable than the wild-type ($\Delta G_U = 1.0 \pm 0.14$ kcal·mol^{-1}). On the other hand, the double deletion mutant has a similar stability to the latter ($\Delta G_U = 1.1 \pm 0.14$ kcal·mol^{-1}) (Table 1). These deletions negatively impact the nucleic acid-binding affinity of *Bc*Csp, as analyzed using fluorescence quenching experiments using dT_6 as the binder (Fig. 1B). The estimated K_D for *Bc*CspΔ36 is ninefold higher than the wild-type, whereas the Δ37 and Δ36-37 deletions lead to a 40-fold increase (Table 1). These differences in affinity may arise due to the effect of the residue deletions on F30, F38, and K39, which are also located in the hinge loop and participate in ligand binding [7]. Moreover, we believe that the deletion present in *Bc*CspΔ36 has a lower effect due to the conservation of G37 and thus providing higher flexibility for correct positioning of contiguous residues involved in ligand binding. Surprisingly, SEC elution profiles after injection of a 2 mM protein sample of *Bc*CspΔ36-37 revealed the presence of two different species with apparent molecular masses of 10.9 and 5.1 kDa, whereas both wild-type *Bc*Csp and the single deletion mutants showed a unique monodisperse species with an estimated molecular mass of 5.1 kDa under the same conditions (Fig. 1C). The difference in the apparent molecular masses between the *Bc*CspΔ36-37 species is consistent with dimer formation. In addition, the dimeric form of *Bc*CspΔ36-37 has a fluorescence quenching behavior that varies almost linearly with dT_6 (data not shown), which might be a consequence of changes in the oligomerization state upon incubation with nucleic acids, as we will show later. Moreover, this dimerization cannot be completely explained by reducing the flexibility of the hinge region as control mutants of E36 and G37 by proline, a strategy usually conceived for modulating domain swapping that reduces the degrees of freedom of the polypeptide chain [39], do not form dimers and exhibit only slight changes in protein stability (Fig. 2). Therefore, additional structural rearrangements may be involved in the dimerization of *Bc*CspΔ36-37.

To gain further insight into this dimer, we analyzed the monomer–dimer equilibrium of *Bc*Csp using analytical SEC, injecting protein concentrations ranging from 50 to 250 μM after incubation for 12 h at three different temperatures between 37 and 50 °C. Our results show an excellent linear correlation between $[M^2]$ and [D] for all assayed temperatures (Fig. 1D). Moreover, van't Hoff analysis of the temperature dependence of the K_D values estimated from the slope of these linear correlations (Fig. 1E) indicated that dimer formation is an enthalpy-driven process by contributing 90% of the total free energy change upon dissociation (Table 2).

Altogether, our results show that deletion of hinge loop residues E36 and G37 allowed dimerization of *Bc*Csp, with no impact on folding stability but compromising the nucleic acid-binding affinity to various extents depending on the deletion. We, therefore, aimed to obtain the crystal structure of this dimeric form of the double deletion mutant in order to determine whether it is generated by domain-swapping interactions or not.

Crystal structure of *Bc*CspΔ36-37 reveals a dimer formed via its ligand-binding surface

High-resolution (< 2.0 Å) crystal structures of Csps from several psychrophilic, mesophilic, and thermophilic organisms including *Bc*Csp have been solved to date [7,9,18,26,30,32], prompting us to attempt the crystallization of the Δ36-37 mutant.

We successfully crystallized and solved the structure of *Bc*CspΔ36-37 starting from the dimeric species previously isolated by SEC (Fig. 3). The crystals belong

Table 1. Thermodynamic parameters of wild-type *Bc*Csp and its single and double-deletion mutants.

	Thermostability			Ligand binding	
Protein	T_m	$\Delta H_{70\ °C}$ [kcal·mol^{-1}]	$\Delta G_{70\ °C}$ [kcal·mol^{-1}]	K_D [μM]	$K_D/K_{D,\ wild\text{-}type}$
Wild-type *Bc*Csp	75.9 (±0.1)	65 (±2)	1.0 (±0.14)	0.04	1
*Bc*CspΔ36	79.2 (±0.1)	63 (±1)	1.5 (±0.14)	0.37	9
*Bc*CspΔ37	82.0 (±0.1)	66 (±1)	2.0 (±0.14)	1.56	39
*Bc*CspΔ36-37	76.6 (±0.1)	61 (±2)	1.1 (±0.14)	1.46	36

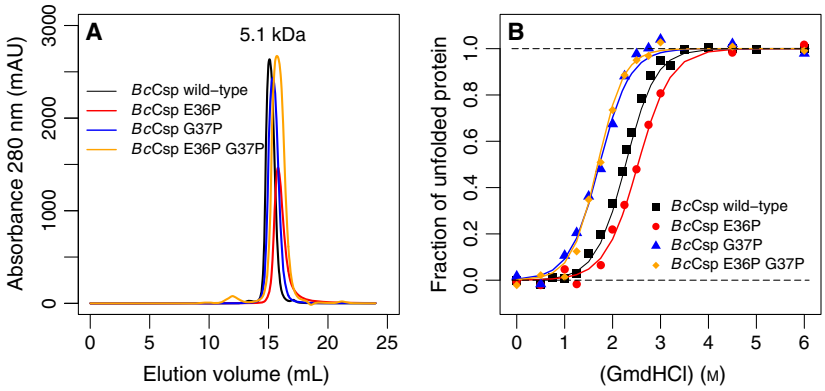

Fig. 2. Reducing flexibility of the hinge loop by introduction of prolines does not lead to dimers but decreases stability. (A) Size exclusion profiles obtained using a Superdex 75 GL 10/300 column followed by absorbance at 280 nm for each mutant at 2 mM protein concentration. The analyzed proteins were wild-type *Bc*Csp (black) and mutants E36P (red), G37P (blue), and E36P-G37P (orange). (B) Reversible protein unfolding induced by guanidine hydrochloride (GdmHCl). The normalized fractions of unfolded protein, based on tryptophan fluorescence emission at 340 nm after excitation at 295 nm, are plotted as a function of concentration of GdmHCl. The curves were fitted to a two-state model (solid lines). The analyzed proteins were wild-type *Bc*Csp (black) and mutants E36P (red), G37P (blue), and E36P-G37P (orange).

Table 2. Thermodynamic analysis of dimer dissociation of *Bc*CspΔ36-37 at several temperatures

T [°C]	$K_D \times 10^{-3}$ (M)	ΔH_D (kcal·mol^{-1})	$T\Delta S_D$ (kcal·mol^{-1})	ΔG_D (kcal·mol^{-1})
50 ± 0.1	1.26 (±0.12)	4.91 (±0.29)	−0.64 (±0.29)	4.27 (±0.41)
42 ± 0.1	1.06 (±0.30)	4.91 (±0.29)	−0.62 (±0.29)	4.28 (±0.41)
37 ± 0.1	0.91 (±0.10)	4.91 (±0.29)	−0.62 (±0.28)	4.29 (±0.40)

to space group $P2_1$ and the structure was solved by molecular replacement to a resolution of 1.8 Å with final R_{work} and R_{free} values of 0.15 and 0.18, respectively (Table 3). The asymmetric unit of the crystal structure of *Bc*CspΔ36-37 contains a dimer (Fig. 3A), consistent with the SEC experiments at high protein concentrations (Fig. 1C). Surprisingly, the crystal dimer is not formed via domain swapping as was observed for the wild-type protein in the presence of hexathymidine [7], but by surface protein–protein interactions between two compact monomers, where the three-dimensional arrangement of secondary structure elements are almost identical to previously solved monomer structures of *Bc*Csp [32] (Fig. 3A). The lack of domain swapping is clear from the observation of the electron density of the hinge loop, whose conformation is consistent with a compact monomer (Fig. 3B). Also, structural superposition of monomeric *Bc*Csp and a single *Bc*CspΔ36-37 subunit shows that these structures are highly similar, only exhibiting significant deviations at the mutated hinge region (Fig. 3C). One of the reasons for the lack of domain swapping may be the need for an extended conformation of the hinge loop to facilitate the intertwining of the two subdomains. The reduction in its length introduced by the two-residue deletion in *Bc*CspΔ36-37 may be fatal to the viability of the domain-swapped dimer. The latter could be further compromised by the lack of E36, which would normally form an electrostatic interaction with K39 in the extended conformation.

Since most Csps are monomeric, we searched the Protein Data Bank for dimeric forms of Csps and compared them with the asymmetric unit as well as with alternative dimers generated by crystal packing. The greatest similarity is found when comparing the dimer of the *Bc*CspΔ36-37 asymmetric unit to the *St*CspE dimer (PDB accession number 3I2Z, Fig. 3C), which is known to be formed both in solution and in the crystal structure through its DNA-binding surface [30]. The only difference between these dimers is the slight rotation of one of the monomers after superimposition of both structures, but they share the formation of protein–protein interactions near the nucleic acid-binding surface (Fig. 3C).

A deeper analysis of the dimer interface of *Bc*CspΔ36-37 was carried out using the GetArea server [44], which allows calculation of each residue's solvent accessible surface area (SASA). By comparing the solvent accessibility of residues in the monomeric and

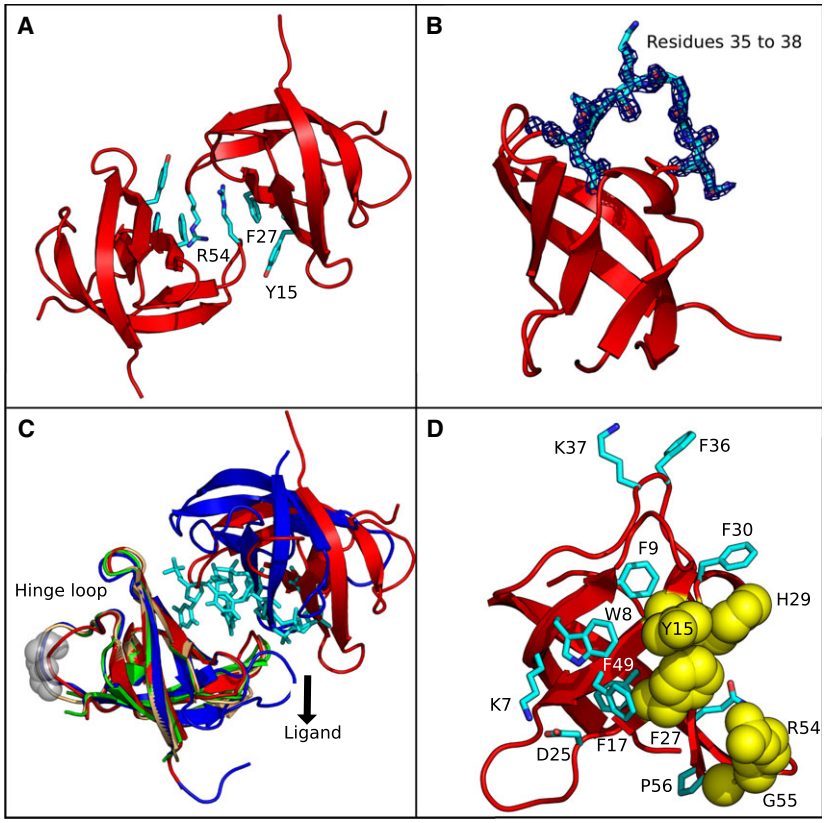

Fig. 3. Crystallographic structure of the nonswapped dimer of *Bc*CspΔ36-37 (A) Cartoon representation of the structure of *Bc*Csp Δ36-37 determined in this work (PDB 5JX4). Each chain is colored in red, whereas residues involved in dimerization are highlighted as sticks. (B) Electron density of the hinge loop (blue mesh contoured at 2.0 σ), demonstrating that each polypeptide chain within the dimer folds into a compact monomer. (C) Structural alignment of wild-type *Bc*Csp (orange), *Bc*Csp Δ36-37 (red), *Bs*CspB in complex with dT_6 (PDB 2ES2, green), showing its ligand in sticks, and *St*CspE (PDB 3I2Z, blue). In addition, the main chain of residues E36 and G37 from wild-type *Bc*Csp are shown as gray spheres (D) Residues involved in dimerization. A closed monomer is shown in red with all residues involved in ligand binding labeled and highlighted using sticks and yellow balls. The latter corresponds to residues occluded in the dimeric structure according to SASA estimations using GetArea.

dimeric states, we could determine which residues become protected upon dimerization and thus constitute the dimer interface. These residues were Y15, F27, G55, H29, and R54 (R56 in wild-type *Bc*Csp) and their accessibility upon dimerization is reduced by 20–60%. Moreover, the PRISM server [45] predicts interfaces formed by the same residues (data not shown). Strikingly, most of these residues are of major functional relevance for *Bc*Csp as they belong to the consensus nucleic acid-binding motifs RNP1 and RNP2 [15]. To provide experimental evidence that the dimer interface does in fact overlap with the nucleic acid-binding surface, we performed irreversible thermal denaturation experiments of the monomer and dimer species of *Bc*CspΔ36-37 with or without dT_6 as a model nucleic acid ligand. Dimeric *Bc*CspΔ36-37 exhibited two thermal transitions in the absence of dT_6, the first one near 50 °C and the second transition similar to the single transition observed for the monomeric form, approximately 75 °C (Fig. 4A,B). By increasing the initial protein concentration of the dimeric form, the first transition became more evident during the unfolding process as it would be expected based on the mass action principle, suggesting that this step would correspond to dissociation of *Bc*CspΔ36-37 into isolated monomers (Fig. 4B). This transition was not observed after addition of dT_6, which exhibited only the single transition that overlaps with monomer unfolding, suggesting that the presence of the ligand destabilizes the dimer, generating monomer:dT_6 complexes rather than dimer:dT_6 (Fig. 4A,B). Considering the K_D for dimerization and ligand binding, where the former is 3 orders of magnitude larger than the latter (Table 1 and 2), it is expected that nucleic acid ligands such as dT_6 would compete for the binding surface and ultimately shift the equilibrium toward the

Table 3. Data collection and structure refinement parameters

Data collection	*Bc*Csp
Space group	$P2_1$
Cell dimensions	
(Å) *a*, *b*, *c*	34.0, 44.1, 36.5;
(°) α, β, γ	90.00, 99.28, 90.00
Detector	R-AXIS IV++
X-ray source	MicroMax 007 HF
Wavelength (Å)	1.5418
Resolution range (Å)	18–1.799 (1.863–1.799)
Multiplicity	3.5 (2.5)
*R*meas (%)	5.7 (14.8)
CC (1/2)	99.8 (96.5)
Completeness (%)	95.10 (75.18)
Total reflections	18 342 (1243)
Unique reflections	9538 (736)
I/σ(I)	23.6 (8.8)
Refinement parameters	
Reflections used for refinement	9514
R (%)	15.00
R_{Free}(%)	18.21
No. of protein atoms	1049
No. of water molecules	221
B (Å^2)	
Protein	7.48
Ligands	22.81
Water	16.21
Error estimates	
Coordinate error (Å)	0.13
Phase error (°)	16.60
Ramachandran Plot	
Favored (%)	98.44
Allowed (%)	1.56
Outliers (%)	0.00
All-atom clashcore	7.66
RMSD from ideal geometry	
r.m.s. bond lengths (Å)	0.013
r.m.s. bond angles (°)	1.335
PDB accession number	5JX4

ligand-bound monomer form over the dimeric form of *Bc*CspΔ36-37.

These experiments suggest that *Bc*CspΔ36-37 forms an unusual dimer both in the crystal structure and in solution via its nucleic acid-binding surface, which is similar to that seen for *St*CspE. We, therefore, sought to gain further insights into the structure and conformational dynamics of this dimer in solution using HDXMS.

Protective effect of dimer formation on the conformational dynamics of *Bc*Csp

We used HXDMS to compare the deuterium incorporation into wild-type *Bc*Csp, with both the monomeric and dimeric *Bc*CspΔ36-37 (Fig. 5). Amide HDX probes solvent accessibility and protein dynamics and the use of pepsin digestion followed by analytical chromatography and mass spectrometry enable the localization of those parts of the protein more susceptible to exchange [46]. We hoped that the HDXMS on the different forms of the protein in solution would help to determine whether the dimer in solution is consistent with the solved crystal structure [47]. Furthermore, as the aromatic residues of the RNP1 and RNP2 motifs have also been shown to have an important role in protein stability [48], it is interesting to assess the conformational dynamics of dimeric *Bc*CspΔ36-37 in comparison with its monomeric counterparts through this experimental strategy.

The methodology used consisted of allowing samples to exchange for 10 min at 37 °C in deuterated buffer and then quenched, pepsin digested, and analyzed by mass spectrometry. This resulted in a total of 22 peptides from which 17 were equivalent between wild-type and mutant *Bc*Csp and considered for analysis. These covered the whole *Bc*CspΔ36-37 sequence (Fig. 5), which was important for further analyses.

We correlated the relative deuterium uptake from our HDXMS with backbone SASA obtained by GetArea [44], calculations that were previously shown to correlate well when calculated on a solved structure [49]. Therefore, we first compared the extent of exchange of freshly SEC-isolated wild-type *Bc*Csp to the backbone SASA for the wild-type monomeric *Bc*Csp structure (PDB accession number 1C9O), resulting in a good correlation ($r^2 = 0.75$, Fig. 5A), indicating that the deuterium uptake profile of the wild-type protein is consistent with the native structure. We next obtained the extent of deuterium exchange for freshly SEC-isolated dimeric *Bc*CspΔ36-37 and determined how good was its correlation with the backbone SASA obtained from the dimeric structure solved in the present work (PDB accession number 5JX4). Correlations were also determined for two alternative models of the wild-type *Bc*Csp (PDB accession numbers 1C9O and 2HAX) once residues E36 and G37 had been excluded. 1C9O corresponds to a monomer while 2HAX is a domain-swapped dimer in the presence of dT_6. Our deuterium exchange analysis showed better correlation with the structure of dimeric *Bc*Csp Δ36-37 than with either model for the wild-type protein ($r^2 = 0.78$, Fig. 5B and Table 4). As these observed correlations for the wild-type protein and the double deletion mutant were highly similar, they strongly suggest that our dimer in solution is as equal to its solved crystal structure as the wild-type protein in solution is to its corresponding crystallographic structure. In addition, the domain-swapped dimer

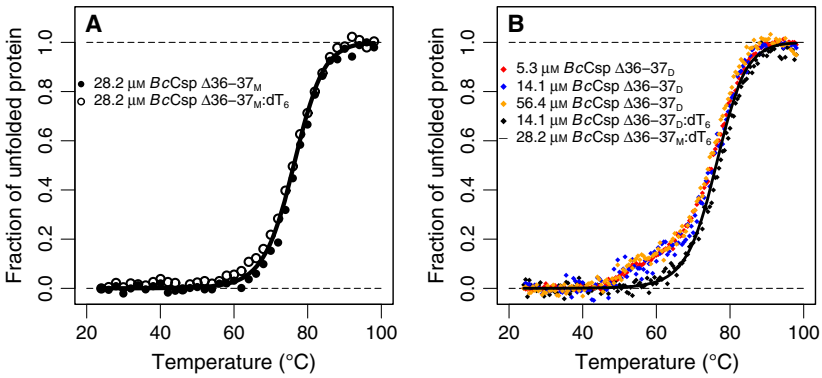

Fig. 4. Dimerization of *Bc*CspΔ36-37 occurs throughout the ligand-binding surface (A) Thermal unfolding of monomeric *Bc*CspΔ36-37 with (empty circle) and without (filled circle) its ligand dT_6. (B) Thermal unfolding of isolated dimeric *Bc*Csp Δ36-37 at several protein concentrations in comparison with unfolding of dimeric *Bc*CspΔ36-37 in the presence of dT_6. The normalized fractions of unfolded protein, based on mean residue molar ellipticity obtained at 222 nm, are plotted as a function of temperature. The solid line represents fitting of monomeric *Bc*CspΔ36-37 without dT_6 to a two-state folding mechanism (Eqns 4, 5).

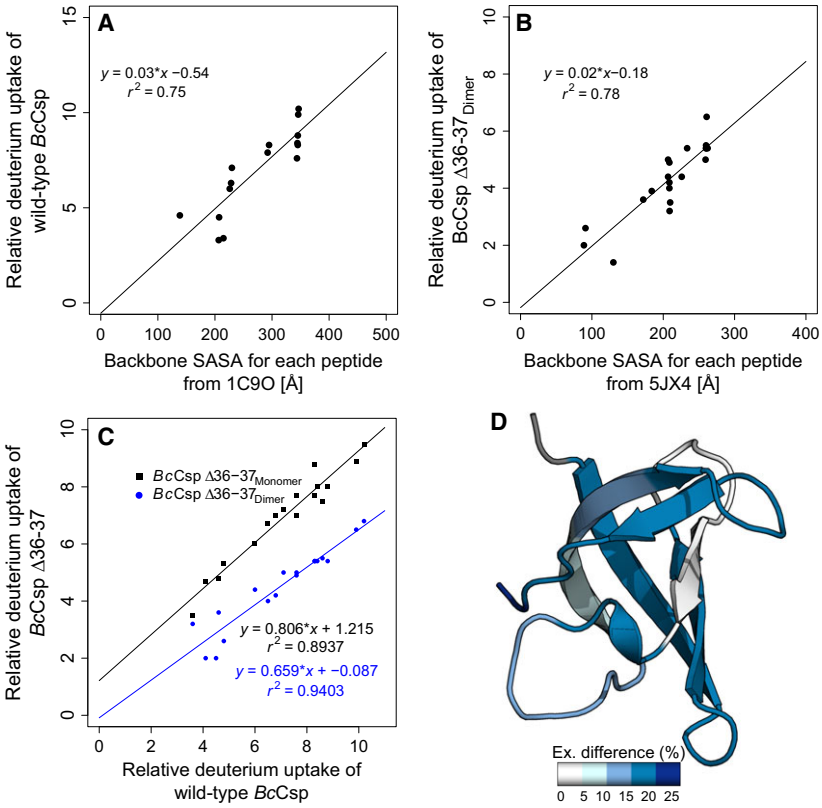

Fig. 5. Protective effect of dimerization of *Bc*CspΔ36-37 in conformational dynamics. (A) Correlation between relative deuterium uptake of wild-type *Bc*Csp and backbone SASA calculated from the crystal structure of wild-type *Bc*Csp (PDB 1C9O). (B) Correlation between relative deuterium uptake of dimeric *Bc*Csp Δ36-37 and backbone SASA calculated from the crystal structure of *Bc*CspΔ36-37 (PDB 5JX4). The solid lines in A and B represent a linear fit for each dataset (C) Correlation of relative deuterium uptake of monomeric (black) and dimeric *Bc*CspΔ36-37 (blue) against wild-type *Bc*Csp, showing that dimerization leads to a decrease in HDX. Only equivalent peptides between wild-type and mutant *Bc*Csp were considered in the analysis (17 out of the 22 peptides), with each point representing a different peptide. The relative deuterium uptake was determined for each peptide obtained after pepsin digestion. The solid lines represent the linear fit for each dataset. (D) Cartoon representation of the difference in relative deuterium uptake upon dimerization of *Bc*CspΔ36-37.

(PDB accession number 2HAX) of *Bc*Csp did not exhibit a good correlation between total SASA and relative deuterium uptake clearly, confirming that the dimer we observe in solution is not domain swapped (Table 4).

We then compared the extent of exchange of the *Bc*CspΔ36-37 dimer against both the monomeric form of the double deletion mutant as well as the monomeric form of wild-type *Bc*Csp. Typically these comparisons between the free subunits and the protein complex allow the identification of the protein–protein interface due to its reduced solvent accessibility upon dimerization, whereas all other protein regions show no significant differences in HDX [50]. Strikingly, dimer formation in the *Bc*CspΔ36-37 mutant leads to a dramatic decrease in the conformational dynamics of *Bc*CspΔ36-37 throughout the whole protein structure rather than exhibiting a localized decrease in the protein-protein interface, as ascertained by the reduced extent of deuterium incorporation in comparison with the monomeric species of wild-type and mutant *Bc*Csp (Fig. 5C,D). Therefore, our results suggest that dimer formation leads to an unexpected protective effect over the dynamic properties of the whole protein in solution.

In order to better understand this phenomenon, molecular dynamics simulations were performed using coarse-grained models for both the monomeric and dimeric species, and their predicted extents of deuterium exchange per residue were compared. The per-residue ln(Pf) were calculated using the native contact criteria established by Craig *et al.* [51], in which a given residue in the native state is considered to be unavailable for exchange when the number of native contacts it establishes is higher than 1. The differences between ln(Pf) of the dimeric and monomeric species of *Bc*CspΔ36-37 were determined at the same T/T_F temperature. We opted for this comparison instead of comparing directly to the HDXMS results due to (a) the coarse resolution of the HDXMS results (average peptide size ~13) in contrast to the per-residue resolution of the SASA analysis; and (b) our previous validation that the structures of wild-type *Bc*Csp and dimeric *Bc*CspΔ36-37 correlate well with the observed behaviors of the monomeric and dimeric species in solution, respectively (Fig. 5). These differences in protection per residue qualitatively follow the same trend observed when comparing the SASA between these species using GetArea (Fig. 6), demonstrating that the differences are not localized in a particular area, but on several residues throughout the whole protein rather than just the protein–protein interface, suggesting that these observations are reproducible based on the geometry and interactions of the native state seen for the crystallographic dimer. In fact, the regions that showed no differences between dimer and monomer (residues 8–12, 20–24 and residues 30–40) are categorized as loops and they are expected to be equally unstructured and dynamic in both aggregation states. Altogether, our HDXMS experiments and structure-based simulations indicate that dimer formation in *Bc*CspΔ36-37 is qualitatively similar in arrangement to that seen for wild-type *St*CspE in crystal structures, and leads to reduced flexibility throughout the whole protein.

Table 4. Extent of correlation between HDXMS results for *Bc*Csp variants and SASA calculated from crystal structures.

HDXMS-analyzed protein	PDB accession number	r^2
Wild-type *Bc*Csp	1C9O	0.75
Wild-type *Bc*Csp	2HAX	0.37
Wild-type *Bc*Csp	5JX4	0.67
Dimeric *Bc*CspΔ36-37	1C9O	0.27
Dimeric *Bc*CspΔ36-37	2HAX	0.37
Dimeric *Bc*CspΔ36-37	5JX4	0.78

The dimerization of *Bc*Csp from an evolutionary and functional perspective

In this work, we have observed how deletions can lead to conformational alterations that could potentially modify the aggregation state of proteins. Deleting two residues from the hinge loop connecting the Csp subdomains unexpectedly allowed the protein to form a dimeric structure through its ligand-binding surface rather than by allowing domain swapping. This was unexpected as domain-swapping had been anticipated based on previous reports of hinge residue deletions in other proteins and on what had been seen in a recent crystal structure of *Bc*Csp in complex with dT_6. The emerging dimer of *Bc*CspΔ36-37 resolved herein is the first observed in solution for a mutant of *Bc*Csp and its observation provides interesting insights into how a deletion can modulate the aggregation state of a structure, and even alter its conformational dynamics.

The relative orientation of the monomers in the dimeric structure of *Bc*CspΔ36-37 is somewhat similar to that of the homologous protein CspE from *S. typhimurium* (*St*CspE). We hypothesize that differences in the hinge loops in some of the proteins belonging to the Csp family are crucial for the dimeric state, because the motif GEG in the hinge loop, which was disrupted in *Bc*CspΔ36-37, is widely conserved in the monomeric proteins of the Csp family, whereas in cold shock proteins from *Enterobacteriaceae* such as

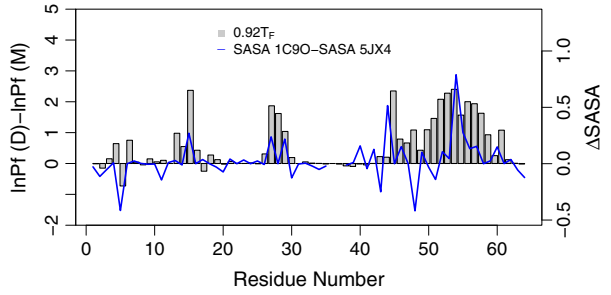

Fig. 6. Structure-based models qualitatively recapitulate the effects of dimerization on the structure of *Bc*CspΔ36-37. The difference in lnPf between dimer and monomer at 0.92T_F is compared to the difference in SASA (blue line) obtained by comparing wild-type monomeric *Bc*Csp (PDB 1C9O) and dimeric *Bc*CspΔ36-37 (PDB 5JX4).

dimeric *St*CspE a TNG motif is present instead (data not shown). We argue that such variations in sequence are important in defining the dimeric state within the Csp protein family and in providing increased stability through a change in aggregation state but its validity remains to be ascertained as the structure and biophysical characterization of many Csps have been focused mainly in their monomeric forms.

Finally, we believe that the reduced conformational dynamics observed as a consequence of dimer formation is a reminiscent of the events that occur due to nucleic acid binding in *Bc*Csp and that are necessary to avoid ligand release and degradation during the cold-shock response. If we consider the fact that localized frustration within proteins is a key feature of protein and ligand-binding interfaces [52], it is plausible that deletions on the hinge loop propagate onto the nucleic acid-binding site by increasing the extent of frustration of this region and also of other regions. In this scenario, the decrease in flexibility of *Bc*Csp upon protein or ligand binding can be rationalized as a way of reducing localized frustration as for several other proteins, but considering the recent characterization of several Csps forming domain-swapped dimers [7,31] and amyloid fibrils [37], it could also constitute a strategy for avoiding the formation of aggregates in conditions where its intracellular protein concentration increases dramatically, as in the case of the cold-shock response.

Materials and Methods

Expression and purification of *Bc*Csp

A codon-optimized gene encoding the thermophilic *Bc*Csp for recombinant expression in *E. coli* was synthesized by GenScript (Piscataway, NJ, USA) and cloned into a modified pET-28a vector that introduces a His-tag followed by a TEV protease cleavage site onto the N-terminal end of the encoded protein. *Bc*Csp residue deletions Δ36, Δ37, and Δ36-37 and proline substitutions E36P, G37P, and E36P-G37P were generated by site-directed mutagenesis.

*Bc*Csps were overexpressed in *E. coli* C41 cells cultured at 25 °C in Terrific Broth containing 35 μg·mL^{-1} kanamycin until reaching an OD_{580} of ~0.8. Protein overexpression was induced by the addition of 0.8 mM isopropyl-β-D-1-thiogalactopyranoside overnight. Cells were harvested by centrifugation, resuspended in 50 mL of buffer A (50 mM Tris-HCl, 500 mM NaCl, 30 mM imidazol, pH 7.8), and lysed by sonication. The crude extract was centrifuged at 36 000 ***g*** and 4 °C for 10 min. The resulting supernatant was heated for 40 min at 60 °C and the denatured proteins were then removed by centrifugation. The supernatant was loaded onto a HisTrap HP column (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) equilibrated with buffer A, and the protein was eluted using a linear gradient of imidazole (30–500 mM). Protein-containing fractions were determined by their absorbance at 280 nm. These were pooled and then dialyzed against buffer B (20 mM Tris-HCl, 5 mM NaCl, pH 8.2) for 16 h. The dialyzed sample was loaded onto a HiTrap Q HP column (GE Healthcare Bio-Sciences) equilibrated with buffer B and the protein eluted using a linear gradient of NaCl (5–1000 mM).

Protein-containing fractions were subjected to proteolytic removal of the N-terminal His-tag using a noncleavable His-tagged TEV protease and then applied again to a Histrap HP column to isolate the cleaved *Bc*Csp. Purity and protein concentration of the pooled, cleaved *Bc*Csp fractions were determined by SDS/PAGE and by absorbance at 280 nm using an extinction coefficient of 6990 M^{-1}·cm^{-1} [53], respectively.

Size exclusion chromatography

Size-exclusion chromatography experiments were carried out on a Water Breeze 1525 HPLC system (Waters Corp., Milford, MA, USA) and protein elution was followed by absorbance at both 220 and 280 nm. Experiments were performed using a Superdex 75 10/300 GL column (GE Healthcare Bio-Sciences) equilibrated with 60 mL of mobile phase, containing 20 mM potassium phosphate pH 8.0 for further CD experiments or 50 mM HEPES, 200 mM NaCl, pH 7.8 for further analytical SEC.

To determine the dissociation equilibrium constant of *Bc*Csp Δ36-37, samples of the monomeric fraction isolated by SEC at several protein concentrations, ranging 50–250 μM in terms of the monomer, were incubated at temperatures between 37–50 °C during 12 h, which proved to be sufficient to reach monomer–dimer equilibrium conditions (data not shown). Samples were centrifuged at 4 °C and 18 000 ***g*** for 10 min to eliminate any particulate matter and kept at 4 °C until injection onto an analytical

TSK-G2000 column (300 mm × 7.8 mm) (Tosoh Bioscience LLC, King of Prussia, PA, USA).

Total concentration was calculated from the absorbance at 280 nm, while fractions of monomer and dimer were quantified from the elution profiles by integrating the peak areas for each species weighted by either dimeric or monomeric mass. Finally, they were analyzed according to a simple dimer-to-monomers dissociation model:

$$D \rightleftharpoons 2M \tag{1}$$

$$K_D = [M^2]/[D] \tag{2}$$

$$\ln(K_D) = -\Delta G_D \cdot R^{-1} \cdot T^{-1} = -\Delta H_D \cdot R^{-1} \cdot T^{-1} + \Delta S_D \cdot R^{-1} \tag{3}$$

where K_D is the equilibrium dimer dissociation constant and Eqn (3) is the van't Hoff equation where R represents the gas constant (0.00198 kcal·mol^{-1}·K^{-1}), T is the temperature in Kelvin and ΔG_D, ΔH_D and ΔS_D are the Gibbs free energy change, enthalpy and entropy of dissociation, respectively.

Columns were calibrated using the Bio-Rad gel filtration standard (Bio-Rad Laboratories, Hercules, CA, USA), composed of: vitamin B-12, 1.35 kDa (8.5 Å Rs); horse myoglobin, 17 kDa (19 Å Rs); chicken ovoalbumin, 44 kDa (30.5 Å Rs); bovine γ-globulin, 158 kDa (41,8 Å Rs); and bovine thyroglobulin, 670 kDa (85,0 Å Rs).

Thermal and chemical unfolding

Thermal unfolding of wild-type *Bc*Csp and its hinge deletion mutants after isolation of monodisperse species by SEC was determined by CD on a Jasco J-815 spectropolarimeter (Jasco Inc., Easton, MD, USA) equipped with a Peltier temperature controller PFD-425S. Samples were heated between 4 and 98 °C at a temperature rate of 1 °C·min^{-1}, collecting data at 222 nm with a response time of 16 s, and a sensibility of 100 mdeg. Experiments were performed in duplicate using a protein concentration of 28 μM in terms of the monomer.

The obtained signal for the reversible thermal unfolding of the monomeric species of all *Bc*Csps was fitted to a two-state N ↔ U model [54] to calculate the free energy change upon unfolding using the Gibbs–Helmholtz equation as follows:

$$K = e^{\frac{-\Delta G_U(T)}{RT}} = \left(\frac{y - y_U}{y_N - y}\right) \tag{4}$$

$$\Delta G_U(T) = \Delta H_m \cdot \left(1 - \frac{T}{T_m}\right) - \Delta C_p \cdot \left\{T_m - T + T \cdot \ln\left(\frac{T}{T_m}\right)\right\} \tag{5}$$

where $\Delta G_U(T)$ is the Gibbs free energy of unfolding, T is the absolute temperature in Kelvin, T_m is the melting temperature (i.e., when $\Delta G_U(T) = 0$) in Kelvin, ΔC_p is the change in specific heat upon unfolding and corresponds to 0.95 kcal·mol^{-1}·K^{-1} [32], K is the equilibrium constant for the N ↔ U reaction, R is the universal gas constant that corresponds to 0.00198 kcal·K^{-1}·mol^{-1}, y is the observed signal, and y_N and y_U are the signals of the native and unfolded states, respectively. Since *B. caldolyticus* is a thermophilic organism with optimal growth temperature of ~70 °C, ΔG_U is reported at this temperature [32].

For chemical unfolding/refolding experiments of control proline substitutions on the hinge region under equilibrium conditions, native or chemically unfolded samples of wild-type *Bc*Csp and the different mutants were diluted to a concentration of 5 μM in buffer 50 mM Tris-HCl pH 8.0, 100 mM NaCl containing varying concentrations of guanidine chloride (GdmHCl) between 0 and 6 M at 25 °C. Then, intrinsic fluorescence measurements were performed using an excitation wavelength of 295 nm and collecting its emission between 305 and 450 nm on a Shimadzu RF-5301PC spectrofluorometer (Shimadzu Europa GmbH, Duisburg, Germany). The maximum fluorescence emission for wild-type *Bc*Csp and for each mutant was plotted against GdmHCl concentration and then fitted to a two-state folding model [55]:

$$F = \frac{F_N + F_U e^{\frac{-(\Delta G_U + m[\mathrm{GdmHCl}])}{RT}}}{1 + e^{\frac{-(\Delta G_U + m[\mathrm{GdmHCl}])}{RT}}} \tag{6}$$

where F_N and F_U are the fluorescence signals from the folded and unfolded states, ΔG_U is the free energy of unfolding and m is the m-value.

Ligand-binding assays

Single-stranded hexathymidine (dT_6) binding was monitored by fluorescence quenching on a 3 μM protein sample of either wild-type *Bc*Csp or its hinge deletion mutants, using between 0.1–3.0 times of ligand over the initial protein concentration. Tryptophan fluorescence was excited at 295 nm, and the emission was collected at 340 nm. The obtained fluorescence signal was corrected for sample dilution, inner filter effects and background fluorescence.

The change in fluorescence intensity upon titration with dT_6 was fitted to a classical binding model [16] as follows:

$$A = K_D + [P] + n \cdot [L] \tag{7}$$

$$Q = \frac{Q_{max} \cdot \left(A - \sqrt{A^2 - 4 \cdot n \cdot [P] \cdot [L]}\right)}{2 \cdot [P]} \tag{8}$$

where K_D is the equilibrium dissociation constant for the dT_6·*Bc*Csp interaction, Q is the fluorescence

quenching upon titration, Q_{max} is the maximum quenching upon saturation with dT_6, [P] is the protein concentration (here, 3 μM for all samples), [L] is the ligand concentration, and n is the number of binding sites in the protein.

X-ray crystallography

Crystallization conditions were initially screened using a high-throughput ARI Gryphon crystallization robot (Arts Robbins Instruments LLC, Sunnyvale, CA, USA) and the Crystal Screen HT reagent kit (Hampton Research Corp., Aliso Viejo, CA, USA). *Bc*CspΔ36-37 crystals were obtained using the hanging drop vapor diffusion method within 2 weeks of equilibration at 291 K of a protein solution containing 20 $mg{\cdot}mL^{-1}$ *Bc*CspΔ36-37 in 20 mM sodium phosphate pH 4.6 with a reservoir solution containing 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, 30% v/v poly(ethylene glycol) methyl ether 2000 after mixing equal volumes of both solutions.

Complete diffraction data up to 1.8 Å were collected at 100K on a Rigaku MicroMax 007 HF X-ray generator equipped with a R-AXIS IV++ detector (Rigaku Corp., TK, Japan) using a wavelength of 1.54 Å. Data were indexed and integrated with XDS [56] and scaled with AIMLESS [57]. The crystal structure was solved by molecular replacement with PHASER [58], using the β-strands of the domain-swapped dimer of *Bc*Csp (PDB accession number 2HAX) as a search model. The structure was refined using PHENIX [59] and COOT [60] and the final structure was validated using MOLPROBITY [61]. Data collection and refinement statistics are summarized in Table 3. All structure figures were generated using PyMol [62].

Hydrogen deuterium exchange mass spectrometry

Hydrogen deuterium exchange mass spectrometry experiments were carried out on a Synapt G2Si quadrupole time-of-flight mass spectrometer with HD/X technology (Waters) controlled by a LEAP H/DX PAL liquid-handling autosampler (Leap Technologies, Carrboro, NC, USA) as previously described [63]. Monomer and dimer species of *Bc*CspΔ36-37 were obtained by SEC at 4 °C and immediately used for HDXMS experiments (protein concentration ~5 μM), keeping them at ~1 °C to avoid dimer re-equilibration into monomers, as SEC analysis demonstrated that the dimer does not dissociate into monomers when maintained for 48 h at temperatures below 4 °C (data not shown). Wild-type monomeric *Bc*Csp was also analyzed for comparison.

Exchange reactions in deuterated buffer were allowed for 0–10 min at 37 °C and then quenched for 2 min at 1 °C using a solution containing 1% formic acid, 3M guanidine-HCl, pH 2.66. Samples were then subjected to online pepsin digestion and analytical chromatography and electrosprayed into the mass spectrometer in Mobility-TOF-ESI+ mode to collect H/DX data. Infusion and scanning of leu-enkephalin (m/z = 556.277) every 30.0 s was used for continuous lock mass correction.

Peptide fragments for wild-type *Bc*Csp and the Δ36-37 double deletion mutant were identified and scored using the PLGS 3.0 software (Waters), selecting for analysis those present in at least two independent runs and having a score >7 and a mass accuracy of at least 3 ppm. Subsequently, deuterium uptake by wild-type *Bc*Csp and monomeric and dimeric *Bc*CspΔ36-37 were determined by calculating the shift of the centroids of the mass envelopes of these peptides compared to the undeuterated samples using DynamX 3.0 (Waters).

Structure-based models

The monomeric and dimeric forms of *Bc*CspΔ36-37 were simulated using coarse-grained structure-based models generated using SMOG2 [64]. In these models, each residue consisted of a single bead centered at the Cα and all native residue-residue contacts were given attractive 10–12 Lennard–Jones interactions, whereas all other nonlocal interactions were treated as repulsive [65]. Simulations were run near the folding temperature (T_F = 1.22 in reduced units) for 5×10^8 steps using a time step of 0.0005 τ, collecting energy and configurations every 5000 steps. Simulations of the dimeric form included an additional weak harmonic constraint centered at the center-of-mass distance between subunits in keeping with the high protein concentration required for dimerization to occur, whereas the same constraint but centered at a distance of 10.0 nm, where the dimer is fully dissociated, was employed for simulations of the isolated monomers.

Protection factor were calculated through an accessibility criterion defined based on the number of native contacts formed by a given residue (Q_i) in the native state [51], such that:

$$\mathrm{Pf} = \begin{cases} 0 \text{ (open) if } Q_i = 0 \\ 1 \text{ (closed) if } Q_i > 1 \end{cases} \quad (9)$$

The probability of the open (P_{op}) and closed states (P_{cl}) for each residue in the native state was calculated by two-dimensional WHAM over the total number of native contacts (Q) and the open and closed state regions of Q_i, allowing calculation of Pf as:

$$\ln(\mathrm{Pf}) = \ln\left(1 + \frac{P_{cl}}{P_{op}}\right) \quad (10)$$

The difference in ln(Pf) between the dimeric and monomeric species were calculated for simulations at the same T/T_F ratio.

Acknowledgements

This work was supported by Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt grant 1130510 to JB and 11140601 to CARS). AIC and GV were supported by graduate fellowships from Comisión Nacional de Investigación Científica y Tecnológica (Conicyt 221320447 to AIC and 22121989 to GV). The Waters Synapt G2Si instrument used in these experiments was obtained from NIH 1S10OD016234-01 (to EAK). The high-throughput ARI crystallization robot was funded by Fondo de Equipamiento Científico y Tecnológico (Fondequip EQM120208). Richard Garratt received financial support from CNPq and FAPESP.

We gratefully thank Matías Fuentealba and Ricardo Cabrera (Departamento de Biología, Facultad de Ciencias, Universidad de Chile) for access and technical support in the use of the high-throughput crystallization robot; and Dr. Ana P. U. Araújo and Dr. Humberto M. Pereira (Instituto de Física de São Carlos, Universidade de São Paulo, Brazil) for providing access to the CD instrument and for technical support during crystal diffraction, respectively.

Author contributions

AIC, EAK, RG, CARS, and JB planned experiments; AIC, GV, VC, and CARS performed experiments. AIC, GV, VC, DAL, and CARS analyzed data. AIC, JB, RG, and CARS wrote the manuscript; all authors critically revised the manuscript and approved the final version to be submitted.

Conflict of interests

The authors declare no conflict of interests exists.

References

1 Jones PG, VanBogelen RA & Neidhardt FC (1987) Induction of proteins in response to low temperature in *Escherichia coli*. *J Bacteriol* **169**, 2092–2095.
2 Barria C, Malecki M & Arraiano CM (2013) Bacterial adaptation to cold. *Microbiology (United Kingdom)* **159**, 2437–2443.
3 Graumann P, Schröder K, Schmid R & Marahiel MA (1996) Cold shock stress-induced proteins in *Bacillus subtilis*. *J Bacteriol* **178**, 4611–4619.
4 Gualerzi CO, Giuliodori AM & Pon CL (2003) Transcriptional and post-transcriptional control of cold-shock genes. *J Mol Biol* **331**, 527–539.
5 Yamanaka K, Fang L & Inouye M (1998) The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation. *Mol Microbiol* **27**, 247–255.
6 Perl D, Welker C, Schindler T, Schröder K, Marahiel MA, Jaenicke R & Schmid FX (1998) Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat Struct Biol* **5**, 229–235.
7 Max KEA, Zeeb M, Bienert R, Balbach J & Heinemann U (2007) Common mode of DNA binding to cold shock domains: crystal structure of hexathymidine bound to the domain-swapped form of a major cold shock protein from *Bacillus caldolyticus*. *FEBS J* **274**, 1265–1279.
8 Lee J, Jeong KW, Jin B, Ryu KS, Kim EH, Ahn JH & Kim Y (2013) Structural and dynamic features of cold-shock proteins of *Listeria monocytogenes*, a psychrophilic bacterium. *Biochemistry* **52**, 2492–2504.
9 Schindelin H, Jiang W, Inouye M & Heinemann U (1994) Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc Natl Acad Sci USA* **91**, 5119–5123.
10 Murzin AG (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J* **12**, 861–867.
11 Theobald DL, Mitton-Fry RM & Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32**, 115–133.
12 Arcus V (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr Opin Struct Biol* **12**, 794–801.
13 Wolffe AP (1994) Structural and functional properties of the evolutionarily ancient Y-box family of nucleic acid binding proteins. *BioEssays* **16**, 245–251.
14 Graumann PL & Marahiel MA (1998) A superfamily of proteins that contain the cold-shock domain. *Trends Biochem Sci* **23**, 286–290.
15 Schröder K, Graumann P, Schnuchel A, Holak TA & Marahiel MA (1995) Mutational analysis of the putative nucleic acid-binding surface of the cold-shock domain, CspB, revealed an essential role of aromatic and basic residues in binding of single-stranded DNA containing the Y-box motif. *Mol Microbiol* **16**, 699–708.
16 Lopez MM, Yutani K & Makhatadze GI (1999) Interactions of the major cold shock protein of *Bacillus subtilis* CspB with single-stranded DNA templates of different base composition. *J Biol Chem* **274**, 33601–33608.
17 Lopez MM, Yutani K & Makhatadze GI (2001) Interactions of the cold shock protein CspB from *Bacillus subtilis* with single-stranded DNA. Importance of the T base content and position within the template. *J Biol Chem* **276**, 15511–15518.
18 Sachs R, Max KEA, Heinemann U& Balbach J (2012) RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution. *RNA* **18**, 65–76.

19 Jiang W, Hou Y & Inouye M (1997) CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. *J Biol Chem* **272**, 196–202.
20 Bae W, Xia B, Inouye M & Severinov K (2000) *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc Natl Acad Sci USA* **97**, 7784–7789.
21 Schindler T, Herrler M, Marahiel MA & Schmid FX (1995) Extremely rapid protein folding in the absence of intermediates. *Nat Struct Biol* **2**, 663–673.
22 Newkirk K, Feng W, Jiang W, Tejero R, Emerson SD, Inouye M & Montelione GT (1994) Solution NMR structure of the major cold shock protein (CspA) from *Escherichia coli*: identification of a binding epitope for DNA. *Proc Natl Acad Sci USA* **91**, 5114–5118.
23 Feng W, Tejero R, Zimmerman DE, Inouye M & Montelione GT (1998) Solution NMR structure and backbone dynamics of the major cold-shock protein (CspA) from *Escherichia coli*: evidence for conformational dynamics in the single-stranded RNA-binding site. *Biochemistry* **37**, 10881–10896.
24 Schnuchel A, Wiltscheck R, Czisch M, Herrler M, Willimsky G, Graumann P, Marahiel MA & Holak TA (1993) Structure in solution of the major cold-shock protein from *Bacillus subtilis*. *Nature* **364**, 169–171.
25 Schindelin H, Marahiel MA & Heinemann U (1993) Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. *Nature* **364**, 164–168.
26 Max KEA, Zeeb M, Bienert R, Balbach J & Heinemann U (2006) T-rich DNA single strands bind to a preformed site on the bacterial cold shock protein Bs-CspB. *J Mol Biol* **360**, 702–714.
27 Zeeb M, Max KEA, Weininger U, Löw C, Sticht H & Balbach J (2006) Recognition of T-rich single-stranded DNA by the cold shock protein Bs-CspB in solution. *Nucleic Acids Res* **34**, 4561–4571.
28 Kremer W, Schuler B, Harrieder S, Geyer M, Gronwald W, Welker C, Jaenicke R & Kalbitzer HR (2001) Solution NMR structure of the cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima*. *Eur J Biochem* **268**, 2527–2539.
29 Jung A (2004) High-temperature solution NMR structure of TmCsp. *Protein Sci* **13**, 342–350.
30 Morgan HP, Wear MA, McNae I, Gallagher MP & Walkinshaw MD (2009) Crystallization and X-ray structure of cold-shock protein E from Salmonella typhimurium. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 1240–1245.
31 Ren J, Nettleship JE, Sainsbury S, Saunders NJ & Owens RJ (2008) Structure of the cold-shock domain protein from Neisseria meningitidis reveals a strand-exchanged dimer. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 247–251.
32 Mueller U, Perl D, Schmid FX & Heinemann U (2000) Thermal stability and atomic-resolution crystal structure of the *Bacillus caldolyticus* cold shock protein. *J Mol Biol* **297**, 975–988.
33 Perl D, Mueller U, Heinemann U & Schmid FX (2000) Two exposed amino acid residues confer thermostability on a cold shock protein. *Nat Struct Biol* **7**, 380–383.
34 Agrawal V & Kishan RK (2001) Functional evolution of two subtly different (similar) folds. *BMC Struct Biol* **1**, 5.
35 Kaneko T (2008) The SH3 domain- a family of versatile peptide- and protein-recognition module. *Front Biosci* **13**, 4938–4952.
36 Makhatadze GI & Marahiel MA (1994) Effect of pH and phosphate ions on self-association properties of the major cold-shock protein from *Bacillus subtilis*. *Protein Sci* **3**, 2144–2147.
37 Alexandrescu AT & Rathgeb-Szabo K (1999) An NMR investigation of solution aggregation reactions preceding the misassembly of acid-denatured cold shock protein A into fibrils. *J Mol Biol* **291**, 1191–1206.
38 Sambashivan S, Liu Y, Sawaya MR, Gingery M & Eisenberg D (2005) Amyloid-like fibrils of ribonuclease A with three-dimensional domain-swapped and native-like structure. *Nature* **437**, 266–269.
39 Rousseau F, Schymkowitz JWH, Wilkinson HR & Itzhaki LS (2001) Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc Natl Acad Sci USA* **98**, 5596–5601.
40 Kelley BS, Leng CC & Bewley CA (2002) Engineering an obligate domain-swapped dimer of cyanovirin-n with enhanced anti-hiv activity. *J Am Chem Soc* **124**, 3210–3211.
41 Perl D & Schmid FX (2001) Electrostatic stabilization of a thermophilic cold shock protein. *J Mol Biol* **313**, 343–357.
42 Delbrück H, Mueller U, Perl D, Schmid FX & Heinemann U (2001) Crystal structures of mutant forms of the *Bacillus caldolyticus* cold shock protein differing in thermal stability. *J Mol Biol* **313**, 359–369.
43 Barrientos LG, Louis JM, Botos I, Mori T, Han Z, O'Keefe BR, Boyd MR, Wlodawer A & Gronenborn AM (2002) The domain-swapped dimer of cyanovirin-N is in a metastable folded state: reconciliation of X-ray and NMR structures. *Structure* **10**, 673–686.
44 Fraczkiewicz R & Braun W (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem* **19**, 319–333.
45 Baspinar A, Cukuroglu E, Nussinov R, Keskin O & Gursoy A (2014) PRISM: A web server and repository for prediction of protein-protein interactions and

modeling their 3D complexes. *Nucleic Acids Res* **42**, W285–W289.
46 Balasubramaniam D & Komives EA (2013) Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim Biophys Acta* **1834**, 1202–1209.
47 Ramírez-Sarmiento CA, Baez M, Wilson CAM, Babul J, Komives EA & Guixé V (2013) Observation of solvent penetration during cold denaturation of *E. coli* phosphofructokinase-2. *Biophys J* **104**, 2254–2263.
48 Schindler T, Perl D, Graumann P, Sieber V, Marahiel MA & Schmid FX (1998) Surface-exposed phenylalanines in the RNP1/RNP2 motif stabilize the cold-shock protein CspB from *Bacillus subtilis*. *Proteins* **30**, 401–406.
49 Truhlar SME, Croy CH, Torpey JW, Koeppe JR & Komives EA (2006) Solvent accessibility of protein surfaces by amide H/2H exchange MALDI-TOF mass spectrometry. *J Am Soc Mass Spectrom* **17**, 1490–1497.
50 Balasubramaniam D, Schiffer J, Parnell J, Mir SP, Amaro RE & Komives EA (2015) How the ankyrin and SOCS box protein, ASB9, binds to creatine kinase. *Biochemistry* **54**, 1673–1680.
51 Craig PO, Lätzer J, Weinkam P, Hoffman RMB, Ferreiro DU, Komives EA & Wolynes PG (2011) Prediction of native-state hydrogen exchange from perfectly funneled energy landscapes. *J Am Chem Soc* **133**, 17463–17472.
52 Ferreiro DU, Hegler JA, Komives EA & Wolynes PG (2007) Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci USA* **104**, 19819–19824.
53 Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD & Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In The Proteomics Protocols Handbook (Walker JM, ed.), pp. 571–607. Humana Press, Totowa, NJ.
54 Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc* **1**, 2527–2535.
55 Santoro MM & Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068.
56 Kabsch W (2010) XDS. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 125–132.
57 Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr* **69**, 1204–1214.
58 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
59 Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 213–221.
60 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 486–501.
61 Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 12–21.
62 The PyMOL Molecular Graphics System, Version 1.7 Schrödinger LLC, http://www.pymol.org.
63 Medina E, Córdova C, Villalobos P, Reyes J, Komives EA, Ramírez-Sarmiento CA & Babul J (2016) Three-dimensional domain swapping changes the folding mechanism of the forkhead domain of FoxP1. *Biophys J* **110**, 2349–2360.
64 Noel JK, Levi M, Raghunathan M, Lammert H, Hayes RL, Onuchic JN & Whitford PC (2016) SMOG 2: A versatile software package for generating structure-based models. *PLoS Comput Biol* **12**, e1004794.
65 Noel JK & Onuchic JN (2012) The many faces of structure-based potentials: from protein folding landscapes to structural characterization of complex biomolecules. In *Computational Modeling of Biological Systems* (Dokholyan NV, ed.), pp. 31–54. Springer, New York, NY, USA.