

# Structural Basis of $\alpha$ -Catenin Recognition by EspB from Enterohaemorrhagic *E. coli* Based on Hybrid Strategy Using Low-Resolution Structural and Protein Dissection

Mitsuhide Hamaguchi<sup>1,2</sup>, Hironari Kamikubo<sup>3</sup>, Kayo N. Suzuki<sup>2</sup>, Yoshihisa Hagihara<sup>4</sup>, Itaru Yanagihara<sup>2</sup>, Ikuhiro Sakata<sup>1</sup>, Mikio Kataoka<sup>3</sup>, Daizo Hamada<sup>5\*</sup>

**1** Department of Emergency Critical Care Medicine, School of Medicine, Kinki University, Osakasayama, Osaka, Japan, **2** Research Institute, Osaka Medical Center for Maternal and Child Health, Izumi, Japan, **3** Laboratory of Bioenergetics and Biophysics, Nara Institute of Science and Technology (NAIST), Ikoma, Nara, Japan, **4** National Institute of Advanced Industrial Science and Technology (AIST), Ikeda, Osaka, Japan, **5** Division of Structural Biology, Department of Biochemistry and Molecular Biology, Graduate School of Medicine, Kobe University, Chuo-ku, Kobe, Japan

## Abstract

Enterohaemorrhagic *E. coli* (EHEC) induces actin reorganization of host cells by injecting various effectors into host cytosol through type III secretion systems. EspB is the natively partially folded EHEC effector which binds to host  $\alpha$ -catenin to promote the actin bundling. However, its structural basis is poorly understood. Here, we characterize the overall structural properties of EspB based on low-resolution structural data in conjunction with protein dissection strategy. EspB showed a unique thermal response involving cold denaturation in the presence of denaturant according to far-UV circular dichroism (CD). Small angle X-ray scattering revealed the formation of a highly extended structure of EspB comparable to the ideal random coil. Various disorder predictions as well as CD spectra of EspB fragments identified the presence of  $\alpha$ -helical structures around G41 to Q70. The fragment corresponding to this region indicated the thermal response similar to EspB. Moreover, this fragment showed a high affinity to C-terminal vinculin homology domain of  $\alpha$ -catenin. The results clarified the importance of preformed  $\alpha$ -helix of EspB for recognition of  $\alpha$ -catenin.

**Citation:** Hamaguchi M, Kamikubo H, Suzuki KN, Hagihara Y, Yanagihara I, et al. (2013) Structural Basis of  $\alpha$ -Catenin Recognition by EspB from Enterohaemorrhagic *E. coli* Based on Hybrid Strategy Using Low-Resolution Structural and Protein Dissection. PLoS ONE 8(8): e71618. doi:10.1371/journal.pone.0071618

**Editor:** Silvio C. E. Tosatto, Universita' di Padova, Italy

**Received:** March 19, 2013; **Accepted:** July 2, 2013; **Published:** August 14, 2013

**Copyright:** © 2013 Hamaguchi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported in part by Grants-in-Aids for the Global COE program A08 (DH), for Scientific Research on Innovative Area "Molecular Science of Fluctuations toward Biological Functions (DH & MK) from MEXT, Japan and for Scientific Research (C) (DH) from JSPS, and the research grants from The Ichiro Kanehara Foundation for the Promotion of Medical Sciences and Medical Care (DH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: daizo@med.kobe-u.ac.jp

## Introduction

EspB is one of the virulence factors of enterohaemorrhagic *E. coli* (EHEC) [1] that is known to be dependent on a type III secretion system (T3SS). EspB is the multifunctional effector with 312 amino acid residues (Fig. 1) that can bind to various different proteins such as  $\alpha$ -catenin [2],[3],  $\alpha_1$ -antitrypsin [4] and myosin [5] from the host cell, and EspA [6] and EspD [7] from the bacterium itself. These interactions of EspB to a range of target proteins are associated with different events in bacterial infection including pore-formation [7], actin reorganization [2],[3],[8] and inhibition of phagocytosis [5].

The binding site for  $\alpha$ -catenin within EspB is known to be located at the N-terminal region between residues 1 to 98 (Fig. 1) [2]. However, the precise mechanism of  $\alpha$ -catenin recognition by EspB is unclear. Previously, we found that binding of EspB to  $\alpha$ -catenin induces dissociation of  $\alpha$ -catenin from an E-cadherin/ $\beta$ -catenin/ $\alpha$ -catenin triple complex formed at the adherence junction on host cell membrane and enhances the intrinsic ability of  $\alpha$ -catenin to promote bundling of actin filaments under *in vitro* conditions [3]. Interestingly, EspB in aqueous solution has the characteristics of a partially folded protein that consists of  $\alpha$ -helical

secondary structures but only small amount of tertiary contacts if any [9]. This conformational property is similar to that of the molten globule state [10–13] which is the compact partially folded state usually accumulated at the early stage of folding kinetics of globular proteins. However, unlike the general molten globule states, EspB does not show an increase of fluorescence intensity of 8-anilino-1-naphthalene-sulfonate [9] which usually binds to surface exposed hydrophobic clusters present in the molten globule intermediates and increase the fluorescence intensity around 480 nm [14]. Moreover, sequence-based disorder predictions as well as a multiplicity of experimental data further suggested that many T3SS-dependent pathogens assume entirely or partially unfolded structures when dissociated from their binding targets [1]. Interestingly, it has been clarified that numerous pathogens from infectious viruses also known to possess a characteristic of intrinsically disordered proteins [15]. Thus, intrinsic disorder could be a generic property for certain classes of virulence factors from infectious bacteria [1] and viruses [15].

Recently, solution NMR has successfully elucidated the structural details of highly disordered proteins. However, such an approach is hardly applicable to EspB due to the line broadening of NMR signals caused by the slow chemical exchange

```

      1      10      20      30      40      50
EHEC EspB: MNTIDNTQVT MVNSASESTT GASSAVAASA LSIDSSLLTD GKVDICKLML
Prediction: -----EE EEE----- ----EEEE-- ----- --HHHHHHHH

      51      60      70      80      90      100
EHEC EspB: EIQKLLGKMV TLLQDYQOKO LAOSYQIQQA VFESONKAIK EKKAATAAL
Prediction: HHHHHHHHHH HHHHHHHHHH HH--HHHHHH HHHHH--HHH HHH---HHHH

      101     110     120     130     140     150
EHEC EspB: VGGIISSALG ILGSFAAMNN AAKGAGEIAE KASSASSKAA GAASEVANKA
Prediction: HHHHHHHHHH HHHHHHHHHH- -----HHHHH HHHHHHHHHH HHHHHHHHHH

      151     160     170     180     190     200
EHEC EspB: LVKATESVAD VAEEASSAMQ KAMATTTKAA SRASGVADDV AKATDFAEDL
Prediction: HHHH---HHH HHHHHHHHHH HHHHHHHHHH- -----HHHH HHHHHHHHHH

      201     210     220     230     240     250
EHEC EspB: ADAAEKTSRI NKLLNSVDKL TTTTAFVAVT SLAEGTKTLP TTISESVKST
Prediction: HHHHHHHHHH HHHHHHHHHH -----EEEE EHH-----HH HHHHHHHH--

      251     260     270     280     290     300
EHEC EspB: HEVNEQRAKS LENFQQGNLE LYKQDVRRTQ DDITTRLRDI TSAVRDLLEV
Prediction: --HHHHHHHH HHHHH-----H HHHHHHHHHH HHHHHHHHHH HHHHHHHHHH

      301     310     312
EHEC EspB: QNRMGQSGRL AG
Prediction: HHHHH----- --

```

**Figure 1. Sequence of EspB from EHEC and the result of secondary structure prediction by Jpred3** (<http://www.compbio.dundee.ac.uk/www-jpred/>) [49]. The sequence region of  $\alpha$ -catenin binding site previously determined [2] is boxed in red and myosin binding sites assigned from the sequence similarity with EPEC EspB [5] is boxed in blue. Regions predicted as helix or strand are indicated as “H” or “E”, respectively. doi:10.1371/journal.pone.0071618.g001

in partially folded regions [9]. Application of another modern high resolution technique such as X-ray crystallography is also impossible because of the difficulty in crystallization of EspB containing significantly flexible regions in the unbound state to its target proteins. We therefore employed a hybrid approach combining low resolution structural and thermodynamic data of intact form and the fragments of EspB obtained by circular dichroism (CD) and small angle X-ray scattering (SAXS) in conjunction with various disorder-order prediction algorithms based on amino acid sequences. The functional properties of the EspB fragments were also tested by fluorescence anisotropy changes upon binding to  $\alpha$ -catenin. Our results provide a significant insight into the structural basis of the recognition of  $\alpha$ -catenin by EspB.

## Materials and Methods

### Materials

All chemicals used in this study were of analytical grade from Sigma-Aldrich Co. (St. Louis, MO), Nacalai Tesque, Inc. (Kyoto, Japan) or Wako Pure Chemical Industries, Ltd (Osaka, Japan). Synthetic peptides of EspB fragments with an N-terminally attached FITC were purchased from PH Japan (Hiroshima, Japan). EspB was prepared using a combination of Ni<sup>2+</sup>-Sepharose and Resource Q chromatography (GE Healthcare, Milwaukee, WI) as described previously [3],[9]. Recombinant EspB were engineered for expression in *E. coli* BL21(DE3) as N-terminally His<sub>6</sub>-tagged fusion proteins using expression vectors derived from pET28a (Merck Chemicals Ltd., Nottingham, UK). The purified proteins were analyzed without

removal of the N-terminal His<sub>6</sub>-tag sequences.  $\alpha$ -Catenin635–906 with an N-terminal His-tag was purified using Chelating Sepharose Fast Flow medium (GE Healthcare) charged with nickel chloride in 20 mM sodium phosphate at pH 7.4. After equilibrating the column in 20 mM sodium phosphate at pH 7.4 the protein solution was applied and unbound material was washed away using the same buffer. Bound protein was subsequently eluted using a 0–0.5 M imidazole gradient. The eluted fractions were dialyzed against 20 mM sodium acetate at pH 4.0 and applied onto a SP-Sepharose column (GE Healthcare) equilibrated in 20 mM sodium acetate at pH 4.0, and then eluted using a 0–2.0 M sodium chloride gradient. Fractions containing  $\alpha$ -catenin635–906 were pooled and further purified by gel filtration chromatography using a Sephacryl S-200 column (GE Healthcare) equilibrated in 20 mM Tris-HCl (pH 8.0), 0.1 mM EDTA. The protein concentrations of full length EspB and EspB1–176 were determined from UV absorption at 280 nm using an extinction coefficient of 4470 and 2980 M<sup>-1</sup> cm<sup>-1</sup>, respectively, which were calculated using the ProtParam server (<http://www.expasy.org/tools/protparam.html>) based on their amino acid sequences. The extinction coefficient for EspB177–312 calculated using the ProtParam server (1490 M<sup>-1</sup> cm<sup>-1</sup>) was too small to obtain reliable estimations of protein concentration in dilute solution. We therefore employed Bradford’s method [16] to determine the concentration of EspB177–312 using standard curve obtained by full length EspB. The absorption spectra were obtained using a JASCO UV spectrophotometer, V-550 (Jasco Co., Tokyo, Japan).

## CD Spectra

Far-UV CD spectra were monitored using a J-720 spectropolarimeter (Jasco Co.) equipped with a Peltier type thermo-controllable cell holder. A quartz cuvette with a 1 cm pathlength was used for the thermal unfolding experiments and a cell with a 0.1 cm pathlength was used for the analysis of CD spectra of EspB fragments as well as full length EspB. Unless stated otherwise, all data are expressed as mean residue ellipticity. Secondary structure contents of the protein and peptides were analyzed by CDpro [17] using the reference set of SP29.

The thermal unfolding transition was monitored by measuring the ellipticity at 222 nm. The temperature was increased from 10 to 90°C at a heating rate of 1°C min<sup>-1</sup>. The samples cooled to 20°C after the thermal unfolding experiments recovered to the essentially same CD spectra obtained at 20°C before heat treatments, suggesting the unfolding reactions were reversible. The thermal unfolding curves were analyzed by modified method as previously demonstrated [18] assuming the two-state transition between folded and unfolded states based on the following equation,

$$\Delta G = \Delta H_0 \left(1 - \frac{T}{T_0}\right) + \Delta C_p (T - T_0) - T \Delta C_p \ln \left(\frac{T}{T_0}\right) + \Delta G_0 \cdot \frac{T}{T_0},$$

where  $T$  and  $T_0$  are the given and reference temperatures in Kelvin,  $\Delta G$  and  $\Delta G_0$  are the free energy changes upon unfolding of folded state at give temperature  $T$  and  $T_0$ , respectively.  $\Delta H_0$  and  $\Delta C_p$  is the enthalpy and the heat capacity change upon unfolding reaction, respectively. The ellipticities at 222 nm ( $[\theta]_{222}$ ) are expressed as  $[\theta]_{222} = [\theta]_F(1 - f_U) + [\theta]_U f_U$ , where  $[\theta]_F$  and  $[\theta]_U$  are the baselines of folded and unfolded species that are linearly dependent on temperature, and  $f_U$  is the fraction of unfolded state.

Thus,  $\Delta G$  can be expressed as  $\Delta G = -RT \ln \left(\frac{f_U}{1 - f_U}\right)$ , where  $R$  is the gas constant. In practice,  $T_0$  was set to 293.15 K (= 20°C), and all the unfolding curves observed in the presence of different amounts of guanidium hydrochloride (GdnHCl) were simultaneously analyzed by the global-fitting method according to the above equations using the same baselines of  $[\theta]_F$  and  $[\theta]_U$ ,  $\Delta C_p$  and  $\Delta H_0$ . IgorPro ver 6.3 (Wavemetrics, Inc., Portland, Oregon) was used for this fitting analysis. The  $\Delta G_0$  values were obtained for the individual conditions at different concentrations of GdnHCl. In the absence of GdnHCl,  $[\theta]_{222}$  was linearly decreased as in crease in  $T$ , and this linear dependence of  $[\theta]_{222}$  was considered as the baseline,  $[\theta]_F$ . This situation prevented us from obtaining the thermodynamic parameters in the absence of GdnHCl based on the above analysis using the data from the thermal unfolding experiments. We therefore estimated  $\Delta G$  in the absence of GdnHCl ( $\Delta G_{\text{water}}$ ) from the linear dependence of  $\Delta G_0$  as a function of the concentration of GdnHCl ( $[\text{GdnHCl}]$ ), i.e.  $\Delta G_0 = \Delta G_{\text{water}} - m[\text{GdnHCl}]$ , which is generally used for the analysis of denaturant-induced unfolding reaction. The  $m$  value is the measure of cooperativity of denaturant-induced unfolding reaction which is considered to be correlated with the change in solvent accessible surface area upon unfolding reaction.

## SAXS

SAXS measurements were carried out at a BL-10C synchrotron beamline in the Photon Factory, Tsukuba, Japan. Scattering intensity was monitored by R-axis VII (Rigaku, Tokyo, Japan) and circularly averaged intensity was used for further analysis. A cell with a 1 mm pathlength was used and the protein solution was prepared by dialysis against 10 mM 3-morpholinopropane-1-sulfonate (MOPS) at pH 7.0. The temperature in the cell was

maintained at 20°C by circulating temperature-controlled water. Scattering intensities obtained at various concentrations of protein were subtracted from the intensity of the buffer solution without protein using IgorPro ver 6.1 (Wavemetrics, Inc., Portland, Oregon). These values were used to generate Guinier and Kratky plots.  $P_r$  functions were calculated by GNOM [19] using the scattering intensity after reduction of noise by singular value decomposition *via* IgorPro. The  $P_r$  function of phosphotriesterase was calculated using a theoretical scattering curve calculated by CRY SOL [20] based on the atomic coordinates of 1EYW solved by X-ray crystallography [21].

## Fluorescence Anisotropy

Fluorescence spectra were observed using a FP-6500 fluorimeter (Jasco Co., Japan). To observe fluorescence anisotropy, HN32 linear polarizer films (5 cm×5 cm; Polaroid Co., Waltham, MA) were inserted in both excitation and emission lightpaths. The orientations of the polarizers were manually changed to obtain the emission components of  $I_{vv}$ ,  $I_{vh}$ ,  $I_{hv}$  and  $I_{hh}$ .  $I_{vv}$  and  $I_{vh}$  were the vertical and horizontal emission components obtained for the sample excited with vertically polarized light, and  $I_{hv}$  and  $I_{hh}$  were the vertical and horizontal emission components obtained for the sample excited with horizontally polarized light. Anisotropy value,  $r$  is then defined as,

$$r = \frac{(I_{vv} - G \cdot I_{vh})}{I_{vv} + 2 \cdot G \cdot I_{vh}},$$

where  $G$  is defined as  $I_{hv}/I_{hh}$ . The data were processed by IgorPro ver 6.3 (Wavemetrics, Inc.) to calculate the anisotropy values with different samples and to estimate the dissociation constants,  $K_D$  by nonlinear curve-fitting.

## Results

### Cooperativity of the Thermal Unfolding Transition of EspB

Thermal unfolding transition of EspB at pH 7.0 was monitored by the ellipticity at 222 nm (Fig. 2). Usually, globular proteins with rigid secondary and tertiary structures stabilized by hydrophobic interactions show highly cooperative thermal unfolding transitions represented by a sigmoidal curve that are consistent with a two-state model from the well-ordered and rigid native state to the highly unfolded state. However, EspB displays a noncooperative transition from a  $\alpha$ -helical structure to a unfolded state in which the absolute ellipticity at 222 nm decreased almost linearly with increasing temperature (Fig. 2A and circles in Fig. 2B). The lack of cooperativity in the thermal unfolding reaction suggests the lack of highly ordered tertiary contacts stabilized by hydrophobic interactions in EspB. A similar behavior was previously observed for the thermal unfolding of a partially folded or “molten globule” intermediate state of human  $\alpha$ -lactalbumin [22]. Thus, the result is consistent with our previous conclusion suggesting that EspB is a natively partially folded protein [1],[9]. Interestingly, further analysis on the response of EspB against the temperature change by far-UV CD revealed the relatively cooperative cold and heat denaturations in the presence of various amount of GdnHCl (Fig. 2B). This result indicates the positive heat capacity change for the unfolding reaction ( $\Delta C_p$ ) of EspB. We analyzed the transition curves assuming two-state transition by global-fitting procedure according to the method described in Materials and Methods. The estimated parameters of unfolding reactions at various GdnHCl concentrations are summarized in Table 1. In this analysis, the estimated  $\Delta C_p$  was  $0.34 \pm 0.01$  kJ mol<sup>-1</sup> K<sup>-1</sup> or  $81.8 \pm 1.0$  cal

$\text{mol}^{-1} \text{K}^{-1}$ . According to the BPPred server (<http://www.clarke.ch.cam.ac.uk/BPPred.php>) [23],  $\Delta C_p$  of a globular protein with 332 amino acid residues, i.e. equivalent length of EspB with hexahistidin-tag, should be  $5.15 \text{ kcal mol}^{-1} \text{K}^{-1}$ . The experimentally determined  $\Delta C_p$  by thermal unfolding experiment of EspB is therefore extremely smaller than that is expected for the globular protein with the same length of EspB.

These results from the CD spectroscopy suggest that either the molecule is in the relatively ordered molten globule-like structure or there is a small relatively structured core and other regions with less-ordered regions in EspB. However, the fact that the EspB does not increase the fluorescence intensity of ANS [9] supports the latter possibility.

### SAXS Indicates the Extended Conformation of EspB

To gain further insight into the conformational properties of EspB, we performed SAXS experiments for this protein at pH 7.0, 20°C. SAXS provides information about the molecular shape and dimension of a macromolecule under a given solution condition. This particular technique is extremely useful when the protein assumes less-ordered structures such as partially folded or highly unfolded states having conformational heterogeneity or ensemble for which more detailed analysis using X-ray crystallography or solution NMR are nearly impossible [24–27].

The pair distance distribution ( $P_r$ ) function calculated from the scattering intensity of EspB using GNOM program [19] showed an extremely broad distribution with a maximum distance of 250 Å and the optimum position of  $P_r$  at 65 Å (Fig. 3A). For comparison, we calculated  $P_r$  function of phosphotriesterase (364 amino acids) as a typical model of a globular protein whose molecular weight is similar to that of EspB (332 amino acids including hexahistidine-tag) using its atomic coordinates solved by X-ray crystallography (Fig. 3A). Unlike the  $P_r$  function of EspB,

the  $P_r$  function of phosphotriesterase indicates an optimum position of  $P_r$  at 24 Å and the maximum distance is  $\sim 60$  Å.

We also determined the radius of gyration ( $R_g$ ) of EspB according to a Guinier plot ( $\ln(I(Q))$  vs.  $Q^2$ ) as shown in Figure 3B. The slope of this plot at low  $Q$  regions corresponds to  $R_g^2/3$  although the value also depends on the protein concentration. To correct the effect of protein concentration dependency, the  $R_g^2$  values were plotted as a function of protein concentration and the reduced  $R_g$  value was given by linear extrapolation of this plot to zero protein concentration (Fig. 3C). This analysis provided the  $R_g$  value of EspB of  $67.8 \pm 0.1$  Å, which is slightly larger than the expected  $R_g$  value of urea unfolded state (62.0 Å) according to the empirical equation [28],  $R_g = 1.927 \times N^{0.598}$ , where  $N$  is the number of amino acids, i.e. 332 in the case of our recombinant EspB with N-terminal hexahistidine-tag.

Scattering intensity obtained by linear extrapolation of the Guinier plot to zero angle,  $I(0)$  is linearly correlated with the molecular weights of the proteins. Molecular weight of EspB was estimated to be 32,400 by comparison of  $I(0)$  of lysozyme with a molecular weight of 14,300. This value is close to the calculated molecular weight of monomeric hexahistidine-tagged EspB (34,800) used in this study. Thus, EspB assumes a highly extended monomeric structure comparable to the ideal random coil. Nonetheless, EspB contains folded regions with  $\alpha$ -helical secondary structures according to the far-UV CD spectra (Fig. 2) [9].

### $\alpha$ -Helical Structures are Formed at the N-terminal Regions Whereas the C-terminal Half is Highly Unfolded

Although the data from SAXS suggested the highly extended structure of EspB, the protein contains a significant amount of  $\alpha$ -helical structures (Fig. 2) [9]. The most  $\alpha$ -helical structures may be partially folded according to the noncooperative behavior of thermal unfolding reaction (Fig. 2) Thus, the overall structure of EspB should consist of both relatively ordered and less well-organized  $\alpha$ -helical structures. To establish the regions of EspB that adopt  $\alpha$ -helical structures, we performed combinatorial analysis using sequence-based disorder-order prediction algorithms and CD spectra of peptide fragments derived from the EspB sequence.

For disorder probability prediction, we used three different algorithms including PONDR [29–31] (<http://www.pondr.com/>), IUPred [32] (<http://iupred.enzim.hu/>) and POODLE-L [33] (<http://mbs.cbrc.jp/poodle/poodle.html>) (Fig. 4). Intriguingly, all three algorithms predicted that the N-terminal half of EspB has a low disorder probability, i.e. relatively high propensity to form ordered structures. By contrast, the C-terminal half of this protein was predicted to have a high propensity to form a disordered structure. To confirm the results of these predictions, we prepared two fragments of EspB corresponding to the amino acid sequences from M1 to T176 (EspB1–176) and from T177 to G312 (EspB177–312) (Fig. 5A) using an *E. coli* expression system and analyzed their ability to form secondary structures by far-UV CD (Figs. 5B and 5C).

As expected, the far-UV CD spectrum of EspB1–176 showed the minima around 210 and 222 nm which are indicative of the presence of a  $\alpha$ -helical structure, whereas that the spectrum of EspB177–312 was typical of an unfolded protein with a minimum around 200 nm (Fig. 5B). Furthermore, the sum of the CD spectra of EspB1–176 and EspB177–312 almost completely reproduced the spectrum of full length EspB (Fig. 5B, broken line). This observation indicates that the  $\alpha$ -helical structures in the N-terminal region are almost independent from the C-terminal unstructured region. Hence, the overall shape of the molecule

**Table 1.** Thermodynamic parameters of unfolding of EspB and EspB41–70 in the presence of GdnHCl at pH 7.0.

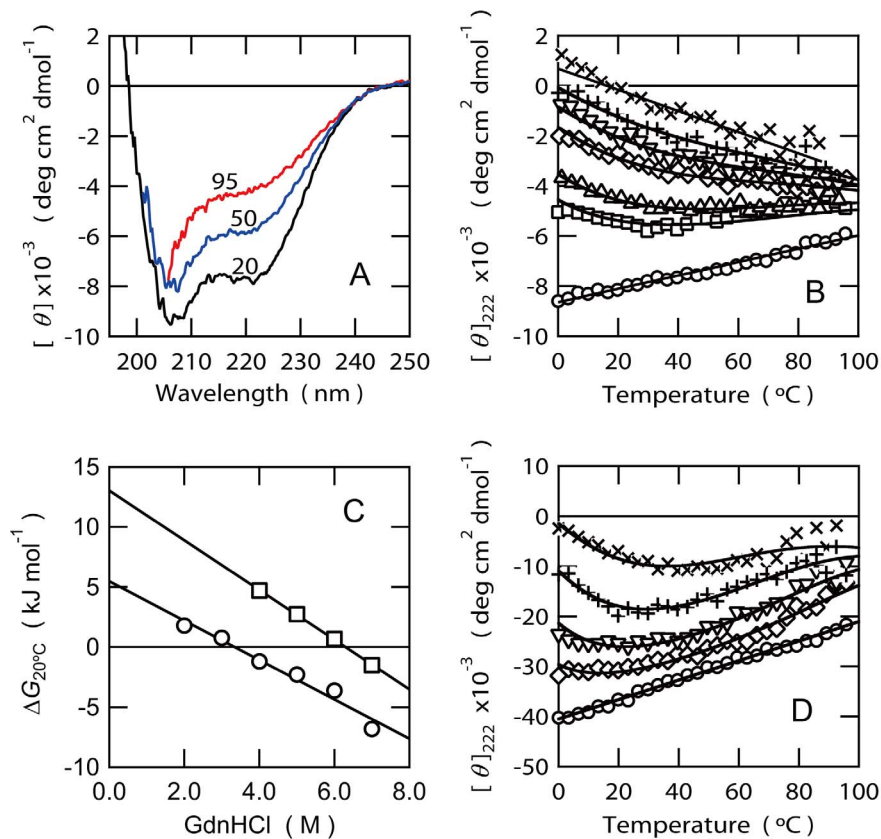
GdnHCl (M)	$\Delta G_{20^\circ\text{C}}$ (kJ mol <sup>-1</sup> )	$\Delta H_{20^\circ\text{C}}$ (kJ mol <sup>-1</sup> )	$\Delta C_p$ (kJ mol <sup>-1</sup> K <sup>-1</sup> )	$m_{\text{GdnHCl}}$ (kJ mol <sup>-1</sup> M <sup>-1</sup> )
EspB				
0.0	5.5±0.7 <sup>a</sup>	n/a <sup>b</sup>	n/a <sup>b</sup>	1.6±0.2 <sup>a</sup>
2.0	1.8±0.1	-9.1±0.1	0.34±0.01	- <sup>b</sup>
3.0	0.8±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>
4.0	-1.2±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>
5.0	-2.3±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>
6.0	-3.6±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>
7.0	-6.8±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>
EspB41–70				
0.0	13.0±0.2 <sup>a</sup>	n/a <sup>b</sup>	n/a <sup>b</sup>	2.1±0.1 <sup>a</sup>
4.0	4.7±0.1	-12.6±0.4	0.87±0.02	n/a <sup>b</sup>
5.0	2.7±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>
6.0	0.7±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>
7.0	-1.5±0.1	SAA <sup>c</sup>	SAA <sup>c</sup>	n/a <sup>b</sup>

<sup>a</sup>Estimated from linear extrapolation of  $\Delta G_{20^\circ\text{C}}$  obtained by the analysis of the data in the presence of GdnHCl as shown in Fig. 1.

<sup>b</sup>Not available.

<sup>c</sup>Same as above.

doi:10.1371/journal.pone.0071618.t001



**Figure 2. Thermal transition of EspB monitored by CD spectra at pH 7.0 in the presence or absence of GdnHCl.** (A) The Far-UV CD spectra of EspB obtained at various temperatures in the absence of GdnHCl. Numbers refer to the temperature in °C. (B) Temperature dependences of the ellipticity at 222 nm ( $[\theta]_{222}$ ) of EspB in the absence of GdnHCl (O) or presence of 2.0 (□), 3.0 (Δ), 4.0 (◇), 5.0 (∇), 6.0 (+) and 7.0 M (X) GdnHCl. Lines are the results of curve-fitting. (C) Linear dependence of  $\Delta G_{20^\circ\text{C}}$  against GdnHCl concentrations to estimate the  $\Delta G_{20^\circ\text{C}}$  in the absence of GdnHCl ( $\Delta G_{\text{water}}$ ) and  $m$  value from the slope of the plots. The  $\Delta G_{20^\circ\text{C}}$  obtained by the thermal transition curves of EspB and EspB41–70 are shown as circles and squares, respectively. (D) Temperature dependences of the ellipticity at 222 nm ( $[\theta]_{222}$ ) of EspB41–70 in the absence (O) or presence of 4.0 (◇), 5.0 (∇), 6.0 (+) and 7.0 M (X) GdnHCl. Lines are the results of curve-fitting.  
doi:10.1371/journal.pone.0071618.g002

comprises a partially folded N-terminal region with a significant amount of  $\alpha$ -helical secondary structure and a largely unstructured C-terminal region. Importantly, EspB1–176 includes the region that was shown to be the  $\alpha$ -catenin binding site [2] (from residue number 1 to 98, see also Fig. 1), suggesting that the  $\alpha$ -helical region of this protein should be, at least partly, involved in the binding site for  $\alpha$ -catenin.

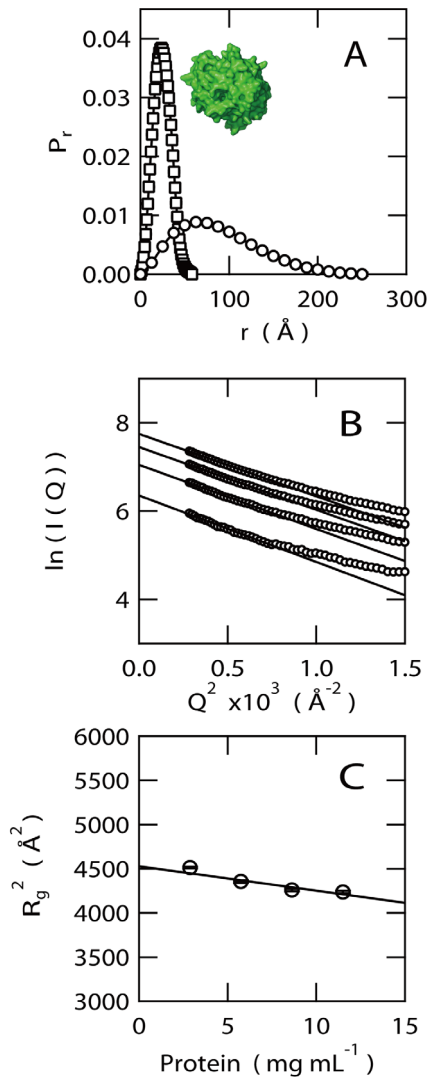
### Structural Properties of $\alpha$ -Catenin Binding Site of EspB by Protein Dissection Analysis

To further refine the structural properties of the regions required for  $\alpha$ -catenin binding in EspB, we prepared several short fragments of EspB, each 20 amino acids in length, covering the first 100 residues from the N-terminus (Fig. 5A). To facilitate the estimation of peptide concentration, all the prepared fragments were modified by FITC through the N-terminal  $\alpha$ -amino group. This modification was also useful in the study of fluorescence anisotropy as described later.

Far-UV CD spectra of the four fragments, covering the region from G41 to E90 (i.e. EspB41–60, EspBN51–70, EspBN61–80, EspBN71–90), indicated the presence of a significant amount of secondary structures (Fig. 6A). By contrast, the other fragments gave spectra corresponding to highly unfolded structures (Fig. 6B). According to the ellipticity around 222 nm, both EspB41–60 and

EspB51–70 have significantly higher  $\alpha$ -helical propensities than those of EspB61–80 and EspB71–90. Interestingly, the CD spectrum of EspB41–60 indicates intensity minima at 222 nm which is greater than that at 208 nm. The ellipticity ratio at 222 vs 208 nm ( $[\theta]_{222}/[\theta]_{208}$ ) larger than 1 is possibly an indication of the formation of either the coiled-coils or other assemblies of helices [34]. Since  $[\theta]_{222}/[\theta]_{208}$  of EspB41–60 is 1.8, the fragment may consist of such rather constraint helices. This and the fact that the SAXS for whole EspB indicates the formation of monomeric structure ruled out the possibility for the formation of oligomeric coiled-coil for EspB41–60.

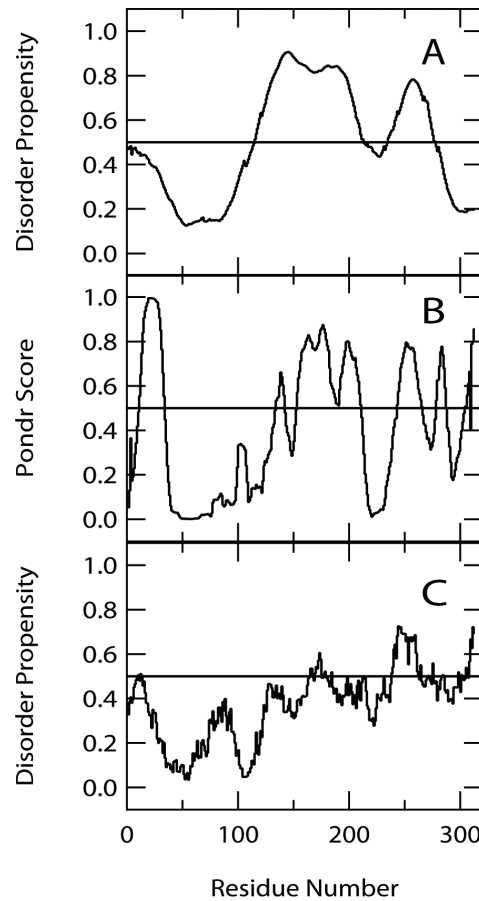
Then, we calculated the secondary structure contents from the CD spectra of the EspB fragments using CDpro software package (Table 2) [17]. This analysis predicted that EspB1–20, EspB11–30, EspB21–40, EspB31–50, EspB81–100 and EspB91–110 possess relatively small amount of secondary structures and nearly 20–30% of  $\beta$ -strand and 10–20% of  $\alpha$ -helix contents but with 55–60% of other structures including unordered structures and turns. EspB41–60 and EspB51–70 exhibited the highest  $\alpha$ -helical contents of 80–90%. Although the spectra for EspB61–80 and EspB71–90 were apparently different from typical  $\alpha$ -helical structure and rather similar to  $\beta$ -rich structure, CDpro suggested that the  $\beta$ -strand contents of these peptides are small (~15%) and  $\alpha$ -helical structures are still dominated (~44%). These results indicates that the  $\alpha$ -helical structures of EspB are probably



**Figure 3. Structural properties of EspB revealed by SAXS.** (A)  $P_r$  function of EspB (circles) calculated from SAXS profile and that of phosphotriesterase (squares) calculated from its atomic coordinates (1EYW) downloaded from Protein Data Bank. The structure of phosphotriesterase is represented as a space-filling model (inset) to illustrate the globular nature of the molecule. This model was produced by PyMol (<http://www.pymol.org>). (B) Guinier plots. Raw data are represented by circles. Lines indicate the results of fitting against linear regions of the raw data. Protein concentrations were 2.0, 5.8, 8.6 and 11.5  $\text{mg mL}^{-1}$  from the bottom to the top in panel B, respectively. (C) Protein concentration dependence of square of  $R_g^2$  calculated from the slope of Guinier plot in panel B. doi:10.1371/journal.pone.0071618.g003

localized around the regions from G41 to E90. For some peptides, the  $\beta$ -strand contents are negative values but the absolute intensities are very small. These values, therefore, can be considered as zero.

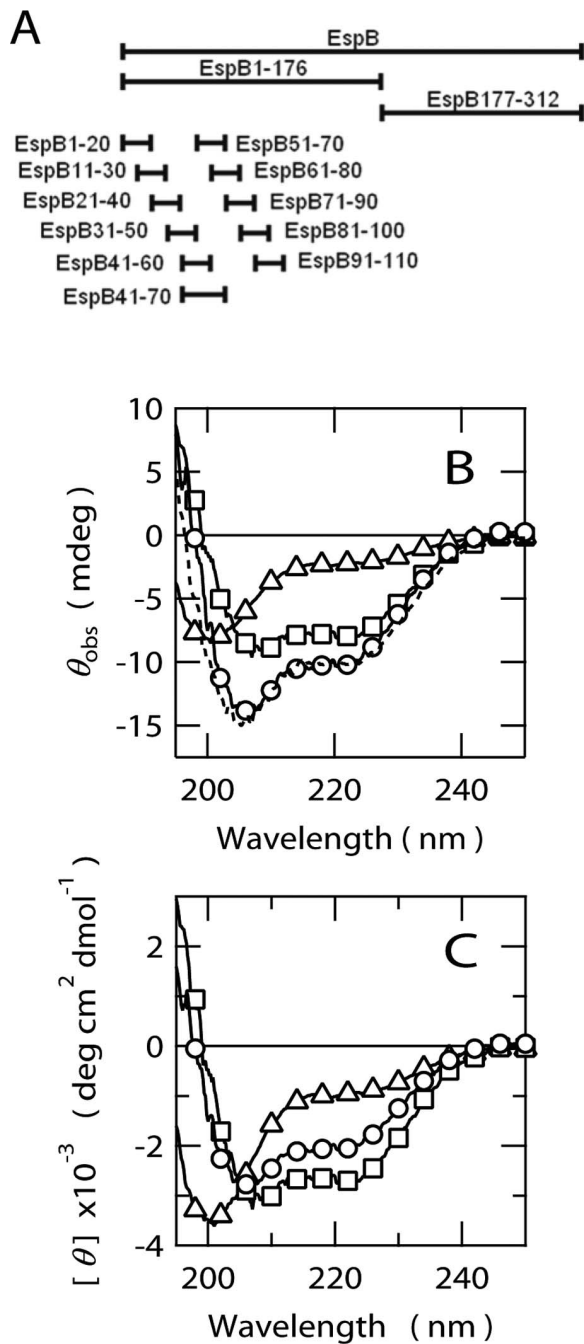
We also compared the results of CDpro analysis with the  $\alpha$ -helical propensities of fragments predicted from the amino acid sequences using AGADIR algorithm [35–39]. The result of AGADIR prediction against the sequence of whole EspB was consistent with the results from the CD spectra of EspB fragments, in which the region around G41 to E90 should have high  $\alpha$ -helical contents. The results of the AGADIR predictions against



**Figure 4. Order-disorder predictions of EspB by using various algorithms.** (A) Poodle-L (<http://mbs.cbrc.jp/poodle/poodle.html>), (B) PONDR (<http://www.pondr.com>) and (C) IUPred (<http://iupred.enzim.hu>). doi:10.1371/journal.pone.0071618.g004

individual peptides qualitatively agree well but quantitatively less well to the estimated  $\alpha$ -helix contents by CDpro.

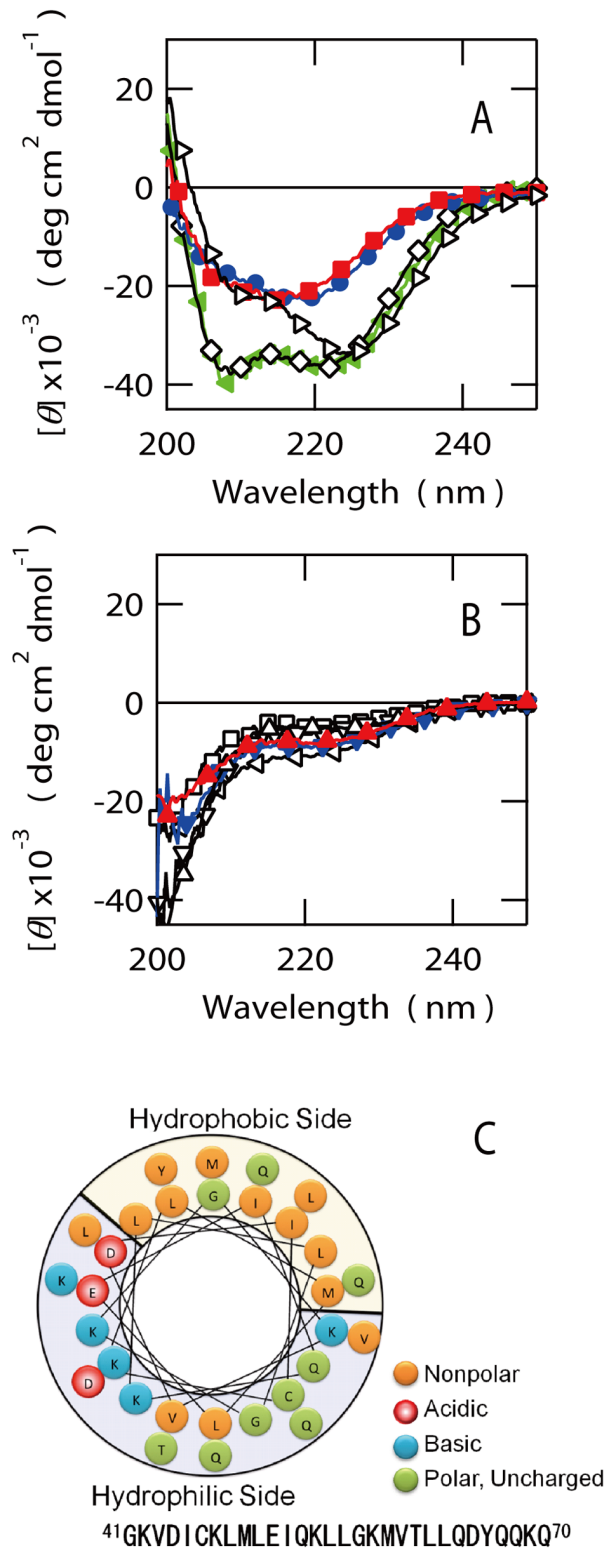
Next, we aimed to identify the fragments that preferentially bind to the C-terminal vinculin homology domain of  $\alpha$ -catenin ( $\alpha$ -catenin635–906), which includes the target recognition region of intact EspB [2], using fluorescence anisotropy of the FITC-labeled EspB fragments (Fig. 7). This experiment clarified that EspB41–60, EspB51–70 and EspB61–80 bind to  $\alpha$ -catenin635–906 (Fig. 7A) although the affinities of EspB41–60 and EspB51–70 were rather higher than that of EspB61–80. Importantly, these two fragments including EspB41–60 and EspB51–70 which indicated the high affinity to  $\alpha$ -catenin635–906 showed the far-UV CD spectra indicative of significant  $\alpha$ -helical propensities (Fig. 6A). However, the  $\alpha$ -helical propensity of EspB61–80 was lower than that of EspB41–60 and EspB51–70. Interestingly, the EspB fragments which showed the CD spectra characteristic of unfolded structures (Fig. 6B) all displayed negligible affinity to  $\alpha$ -catenin635–907 (Fig. 7B). Thus, the data showed a strong relationship between  $\alpha$ -helical propensities of the EspB fragments and their binding abilities to  $\alpha$ -catenin. The  $K_D$  values of EspB41–60 and EspB51–70 were  $23 \pm 2$  and  $12 \pm 3$   $\mu\text{M}$ , respectively. These values are five to ten times higher than that of wild type EspB ( $2.9 \pm 0.3$   $\mu\text{M}$ ), suggesting that neither of the fragments completely mimic the function of intact EspB. The affinity of EspB61–80 was too weak to obtain the reliable  $K_D$  value. An extremely high concentration of  $\alpha$ -catenin635–906 would be required to deter-



**Figure 5. Far-UV CD spectra of EspB fragments.** (A) Schematic representation of the fragments prepared in this study. (B) Raw ellipticity data ( $\theta_{\text{obs}}$ ) obtained by full length EspB (circles), EspB1-176 (squares) and EspB177-312 (triangles). The protein concentration for each protein was 15  $\mu\text{M}$ . Broken line indicates the sum of the spectra of EspB1-176 and EspB177-312. (C) Another representation of spectra shown in panel B expressed as mean residue ellipticity,  $[\theta]$ . doi:10.1371/journal.pone.0071618.g005

mine an accurate  $K_D$  value for this peptide using our experimental conditions.

To further refine the regions of EspB that can reproduce the affinity of intact EspB, we prepared longer fragments corresponding to G41 to E70 (EspB41-70). The far-UV CD spectrum of EspB41-70 was consistent with the formation of  $\alpha$ -helical structures (Fig. 6A). The analysis of CDpro indicated that the  $\alpha$ -



**Figure 6. Structural properties of short fragments of EspB.** (A) Far-UV CD spectra for the peptides assuming  $\alpha$ -helical structures. EspB41-60 ( $\blacktriangleright$ ), EspB51-70 ( $\diamond$ ), EspB61-80 ( $\bullet$ ), EspB71-90 ( $\blacksquare$ ), EspB41-70 ( $\blacktriangleleft$ ). (B) Spectra for the peptides assuming unfolded structures. EspB1-20 ( $\square$ ), EspB11-30 ( $\Delta$ ), EspB21-40 ( $\nabla$ ), EspB31-50 ( $\triangleleft$ ), EspB81-100 ( $\blacktriangle$ ), EspB91-110 ( $\blacktriangledown$ ). (C) Helical wheel representation of EspB41-70, indicating the amphipathic nature of this region. doi:10.1371/journal.pone.0071618.g006

**Table 2.** Secondary structures of EspB and its peptide fragments estimated from the results of far-UV CD or amino acid sequences.

Peptide	$\alpha$ -Helix <sup>a</sup> (%)	$\beta$ -Strand <sup>b</sup> (%)	Turns <sup>c</sup> (%)	Unordered <sup>d</sup> (%)	Helix Propensity <sup>e</sup> (%)	$N_H$ <sup>f</sup>
Whole EspB	22.9±1.2	22.4±1.6	21.8±0.8	34.3±1.1	4.98	76
EspB1–176	35.7±0.9	14.2±1.4	21.1±0.9	28.9±1.1	7.56	70
EspB177–312	9.1±3.6	29.9±1.2	23.5±3.0	37.8±3.7	4.72	14
EspB1–20	9.6±6.6	26.4±3.0	23.0±5.3	40.6±4.5	0.12	2
EspB11–30	11.7±6.3	27.2±6.8	22.6±2.2	39.5±0.9	0.32	2
EspB21–40	12.0±6.4	25.3±5.1	22.5±1.2	38.4±1.1	0.63	2
EspB31–50	24.9±3.3	16.5±4.5	22.5±2.6	35.8±1.8	0.56	5
EspB41–60	81.6±9.7	5.3±5.8	5.6±5.1	7.9±5.0	25.86	16
EspB51–70	89.3±7.0	−2.3±4.9	7.7±8.1	9.6±10.9	0.81	18
EspB61–80	44.0±13.1	15.1±7.7	15.7±3.6	25.7±8.7	3.10	9
EspB71–90	44.1±7.3	14.6±9.9	18.1±3.0	23.4±8.9	0.62	9
EspB81–100	19.1±1.6	22.1±2.9	21.5±2.6	37.4±1.2	5.28	4
EspB91–110	15.1±4.2	21.0±10.3	22.1±1.3	40.8±8.9	0.44	3
EspB41–70	93.9±6.0	−0.5±3.5	8.3±7.6	5.3±7.4	23.48	28

<sup>a</sup>Sum of the contents of regular and distorted  $\alpha$ -helix estimated by CDpro [17] based on the far-UV CD spectra.

<sup>b</sup>Sum of the contents of regular and distorted  $\beta$ -sheet estimated by CDpro [17] based on the far-UV CD spectra.

<sup>c</sup>Contents of turns estimated by CDpro [17] based on the far-UV CD spectra.

<sup>d</sup>Contents of unordered structures estimated by CDpro [17] based on the far-UV CD spectra.

<sup>e</sup> $\alpha$ -Helical propensities calculated by AGADIR algorithm [35–39].

<sup>f</sup>Number of residues involved in  $\alpha$ -helix calculated from the  $\alpha$ -helix content estimated by CDpro [17] and the total length of each polypeptides. Total length of EspB, EspB1–176 and EspB177–312 are 332, 197 and 156 amino acids including additional sequences of N-terminal his-tag.

doi:10.1371/journal.pone.0071618.t002

helical content of EspB41–70 becomes 93.9%, i.e. about 28 amino acid residues may be involved in  $\alpha$ -helical structures in EspB41–70. AGADIR program also predicted that this fragment possess high helix propensity (Table 2). Interestingly, EspB41–70 displayed a unique thermal response similar to that of whole EspB in the presence and absence of GdnHCl (Fig. 2). The estimated  $\Delta C_p$  value for EspB41–70 was  $0.87 \pm 0.02$  kJ mol<sup>−1</sup> K<sup>−1</sup>. This value is rather similar to that of whole EspB ( $0.34 \pm 0.01$  kJ mol<sup>−1</sup> K<sup>−1</sup>). We also estimated  $m$  values for EspB and EspB41–70 upon unfolding reaction by GdnHCl and the  $\Delta G_{\text{water}}$  at 20°C in the absence of GdnHCl from the linear relationship between  $\Delta G_{20^\circ\text{C}}$  and GdnHCl concentration, [GdnHCl], i.e.  $\Delta G_0 = \Delta G_{\text{water}} - m[\text{GdnHCl}]$  (Fig. 1D).  $\Delta G_{\text{water}}$  of EspB ( $5.5 \pm 0.7$  kJ mol<sup>−1</sup>) was smaller than that of EspB41–70 ( $13.0 \pm 0.2$  kJ mol<sup>−1</sup>). Thus, the presence of additional regions at the N- and C-terminal may rather destabilize the relatively rigid  $\alpha$ -helical structure formed around G41 to Q70. The  $m$  values are considered to be correlated with the change in the accessible surface area upon unfolding of a protein. Interestingly, the  $m$  value of EspB ( $1.6 \pm 0.2$  kJ mol<sup>−1</sup> M<sup>−1</sup>) was closely similar to the value of EspB41–70 ( $2.1 \pm 0.1$  kJ mol<sup>−1</sup> M<sup>−1</sup>), suggesting that the structural elements responsible for the rather cooperative thermal response of EspB assume the structure similar to that of EspB41–70. These results suggested that the sequence around G41 to Q70 by itself could form the core of rigid  $\alpha$ -helical structure formed in EspB. Moreover, the  $K_D$  value of EspB41–70 for  $\alpha$ -catenin was  $2.5 \pm 0.3$   $\mu\text{M}$ , which is very close to the value determined for intact EspB ( $2.9 \pm 0.3$   $\mu\text{M}$ ) (Fig. 7A). These results suggest that the sequence around G41 to Q70 of EspB must be the core  $\alpha$ -helical region which is also important for the recognition of  $\alpha$ -catenin. A close analysis of the sequence from G41 to Q70 clarified the possible formation of an amphipathic  $\alpha$ -helix in this region (Fig. 6C) if it assumes extended  $\alpha$ -helical structure. However, it is

unclear if such an extended amphipathic  $\alpha$ -helical structure can be formed under normal solution conditions.

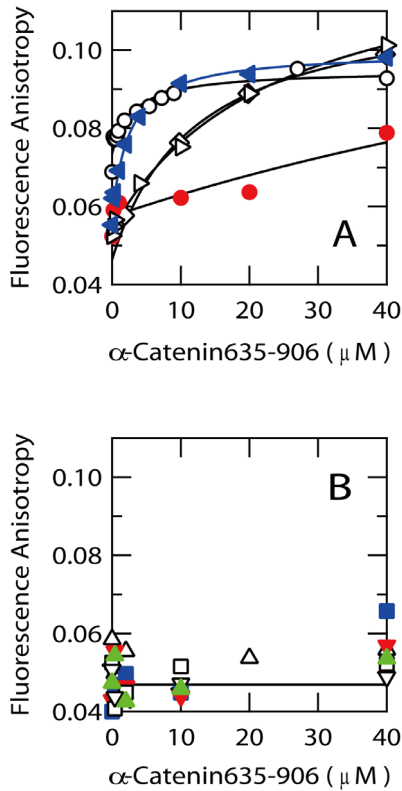
We then analyzed if any additional secondary structures are induced for EspB upon binding to  $\alpha$ -catenin based on far-UV CD spectra. Interestingly, the far-UV CD spectrum for the solutions containing of  $\alpha$ -catenin635–906 and EspB in 1 to 1 ratio coincide very well with the sum of the individual spectra obtained for  $\alpha$ -catenin635–906 and EspB (Fig. 8). This clearly indicates that no additional secondary structures are formed for these proteins upon formation of  $\alpha$ -catenin/EspB complex.

## Discussion

In general, proteins synthesized in biological systems spontaneously folds into well-ordered structures that are unique to their amino acid sequences, thereby conferring distinctive functions. However, our data suggest that EspB possesses highly unique structural properties. Specifically, the N-terminal region around G41 to E70 of EspB has a significant amount of  $\alpha$ -helix contents whilst the C-terminal region is largely unstructured. The analysis of EspB fragments further suggested that the  $\alpha$ -helical region is directly involved in the binding with  $\alpha$ -catenin.

A relatively cooperative thermal response including cold denaturation phenomena observed for EspB and its fragment of EspB41–70 suggests the presence of small but distinctively rigid structures around G41 to Q70 that is stabilized by hydrophobic interactions. However, the contribution of such hydrophobic contacts particularly among the sidechains may be very little since the values of  $\Delta C_p$  were extremely small (Table 1). The length of EspB41–70 is 30 amino acid long and about 28 amino acid residues will be involved in  $\alpha$ -helical structures according to the secondary structure prediction by CDpro. If this estimate is correct, the most parts of this fragment assumes  $\alpha$ -helical structure

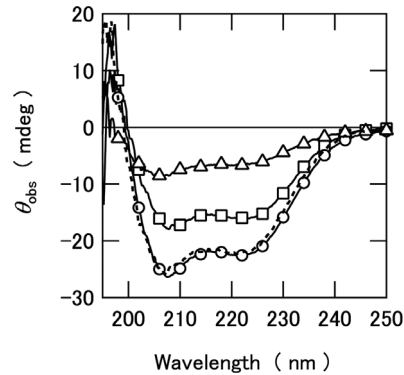




**Figure 7. Affinity of the EspB fragments to  $\alpha$ -catenin635–906 analyzed by fluorescence anisotropy.** Each EspB peptide was modified with a fluorescein moiety at the N-terminus. (A), Titration curves for the peptides bound to  $\alpha$ -catenin. EspB ( $\circ$ ), EspB41–60 ( $\triangleright$ ), EspB51–70 ( $\diamond$ ), EspB61–80 ( $\bullet$ ), EspB41–70 ( $\blacktriangleleft$ ). (B) Titration results obtained for the peptides unable to bind to  $\alpha$ -catenin635–906. EspB1–20 ( $\square$ ), EspB11–30 ( $\Delta$ ), EspB21–40 ( $\nabla$ ), EspB31–50 ( $\triangleleft$ ), EspB71–90 ( $\blacksquare$ ), EspB81–100 ( $\blacktriangle$ ), EspB91–110 ( $\blacktriangledown$ ). doi:10.1371/journal.pone.0071618.g007

and one of the extreme patterns should be an long extended  $\alpha$ -helix. Such an extended  $\alpha$ -helix may possess only little number of sidechain-sidechain contacts and the contribution of hydrophobic effects to the stabilization should be less high. It is of question if such an extended  $\alpha$ -helix structure by a single chain is stabilized by the mechanism same as that of the globular protein even though the helix is sufficiently stable.

Interestingly, secondary structure prediction from the CD spectrum of EspB indicated that about 76 amino acid residues should be involved in  $\alpha$ -helical structures within 332 amino acid sequence of EspB. Therefore, although the thermal unfolding behavior and  $\alpha$ -catenin binding of EspB can be almost successfully explained by the contribution of the sequences from G41 to Q70, some other parts of the protein should be involved in the formation of additional  $\alpha$ -helices. These regions may be rather less organized and behave like the partially folded intermediates of globular proteins, namely the “molten globule” states [10–12], [27]. Interestingly, SAXS data revealed that, although the molten globule states of globular proteins are as compact as the well-defined native states, EspB assumes an extremely extended conformation with the  $R_g$  value closely similar to the ideal random coil structures. In this sense, the less organized regions of EspB should be in the conformational states similar to the “premolten globule” states [13],[40] which are only transiently accumulated at the very early stage of folding of globular proteins. Such an



**Figure 8. Far-UV CD spectrum of the EspB/ $\alpha$ -catenin635–906 complex.** The spectra of EspB ( $\Delta$ ),  $\alpha$ -catenin635–906 ( $\square$ ), and a complex of EspB and  $\alpha$ -catenin635–906 ( $\circ$ ) are shown. Broke line indicates the calculated sum of the spectra of EspB and  $\alpha$ -catenin635–906. The concentrations of both EspB and  $\alpha$ -catenin635–906 were 3.2  $\mu$ M. doi:10.1371/journal.pone.0071618.g008

extended partially folded structure of EspB should be dominantly stabilized by the local interactions formed between amino acids close to each other in the amino acid sequence rather than the nonlocal interactions formed by the residues apart from each other in the sequence.

Taken together, most parts of EspB assume an extended premolten globule-like structures with significant amount of  $\alpha$ -helix and the region of G41 to Q70 forms relatively organized  $\alpha$ -helical structures that are directly responsible for  $\alpha$ -catenin binding. This result is consistent with the observations for various intrinsically disordered proteins that indicate the regions with high propensity to form secondary structures tend to be functionally important [41].

Highly unstructured intrinsically disordered proteins without having any secondary and tertiary structures often show the formation of well-defined structures upon binding to their target molecules in a manner of either “*induced fit*” or “*conformational selection*” [42]. In the former case, the unstructured protein initially binds to its target molecule as an unfolded structure followed by the formation of well-defined three dimensional structure on the target molecule. Whereas, in the latter case, the particular structure that is only transiently formed among the ensemble of the fluctuating structure of the unstructured protein is selectively recognized by the target molecule. The term, “*coupled folding and binding*” is recently used when the secondary structures (usually  $\alpha$ -helix) are formed by induced-fit mechanism [43–44]. No additional secondary structure was induced neither for EspB nor  $\alpha$ -catenin635–906 upon formation of their complex structure according to far-UV CD spectra. This result suggests that the binding of EspB to  $\alpha$ -catenin proceeds by the recognition of the preformed secondary structure in EspB by  $\alpha$ -catenin. This reaction may be classified into an extreme case of “*conformational selection*” or possibly called “*folding before binding*” in which the formation of  $\alpha$ -helix structure that is similar to the helix formed in the protein complex [45]. However, the current data cannot rule out the possibility that additional conformational changes like rearrangements of sidechain orientation or helix curvature without changing total  $\alpha$ -helical content may take place in a manner of *induced fit* reaction even in the case of EspB. This type of reaction modes that is initially driven by conformational selection followed by induced fit has been proposed for the interaction between transactivation domain of tumor suppressor p53 and the nuclear coactivator binding domain of cyclic-AMP response element binding protein

from the result of molecular dynamic simulation [46]. Such a combined mechanism could be the major mechanisms of molecular recognitions particularly for intrinsically disordered proteins having partially folded regions similar to the molten globule or premolten globule states where preformed partially folded structures are involved in binding process. The structural details on the free and complex forms of EspB as well as the kinetic analysis for  $\alpha$ -catenin binding are required to clarify the more realistic mechanism for  $\alpha$ -catenin binding of EspB.

Interestingly,  $\alpha$ -catenin binds to  $\beta$ -catenin through its N-terminal vinculin homology domain whereas EspB binds to the C-terminal vinculin homology domain. This fact suggests that the binding of EspB to the C-terminal region of  $\alpha$ -catenin should induce the conformational change around the N-terminal region of  $\alpha$ -catenin to promote the dissociation of  $\alpha$ -catenin from  $\beta$ -catenin. The analysis of 3D structure of EspB/ $\alpha$ -catenin complex using e.g. X-ray crystallography will clarify the mechanism of the dissociation of  $\alpha$ -catenin from E-cadherin/ $\beta$ -catenin complex by the presence of EspB.

In this paper, we identified that only about 30 amino acid residues in EspB play important role for  $\alpha$ -catenin binding which probably promotes the rearrangement cellular morphology. Then, what is the role of other regions of EspB? As mentioned earlier in this paper, EspB is a multifunctional protein which is involved in pore-formation [7], actin reorganization [2],[3],[8] and inhibition of phagocytosis [5]. The analysis demonstrated here focused on the actin reorganization through binding of  $\alpha$ -catenin. Previous analysis on enteropathogenic *E. coli* (EPEC) indicated that the region from I159 to L218 in EPEC EspB is important for binding to myosin [5]. The regions in EHEC EspB corresponding to this region is relatively well conserved and are assigned to V158 to R109 [5] (Fig. 1). Thus, this region in EHEC EspB can also involved in inhibition of phagocytosis by binding to myosin proteins. The regions responsible for other binding partners of EspB are still unknown, but the regions other than G41–Q70 or V158 to R109 can be also required for recognition of various other proteins. Simultaneous binding of various other proteins using

different binding site can also important for all the phenomena that EspB promotes. Then, why EspB has to assume less-organized extended structures with partially folded  $\alpha$ -helices? The observations of various other effectors that are secreted through the type III secretion systems of other infectious bacteria suggest that they tend to form less-organized structures as we observed for EspB [1]. This fact implicates that the conformational disorder of T3SS effectors are required for effective secretion of these proteins through the narrow central channel of T3SS needles with inner diameters of  $\sim 20$  Å [47]. This size allows the secretion of the globular proteins with  $< 15$  amino acids in length or the hydrodynamic radius ( $R_h$ ) of  $< 10$  Å, according to the empirical correlation of  $R_h = 4.75 \cdot N^{0.29}$  (Å) where  $N$  is the length of a protein [48].

In conclusion, we identified that the functionally important regions of EspB forms  $\alpha$ -helical structures stabilized mainly by local interactions around G41 to Q70 and other regions are less well-organized. As a consequence, the overall structure of EspB is as extended as random coil structures. This result is highly relevant for understanding the mechanism of  $\alpha$ -catenin recognition by EspB. Further analysis for the structural property of the complex between EspB and  $\alpha$ -catenin will provide further detail of the mechanism of  $\alpha$ -catenin binding by EspB. Such information will be also critical for understanding more precise structural details of unbound form of EspB in solution.

## Acknowledgments

We thank M. Ozawa (Kagoshima University, Japan) for providing us with a cDNA clone of  $\alpha$ -catenin.

## Author Contributions

Conceived and designed the experiments: DH MH HK YH MK. Performed the experiments: DH MH KNS HK YH. Analyzed the data: DH HK MK. Contributed reagents/materials/analysis tools: DH MH KNS. Wrote the paper: DH HK MH YH IY IS MK.

## References

- Hamada D, Hamaguchi M, Suzuki KN, Sakata I, Yanagihara I. (2010) Cytoskeleton-modulating effectors of enteropathogenic and enterohemorrhagic *Escherichia coli*: a case for EspB as an intrinsically less-ordered effector. *FEBS J* 277: 2409–2415.
- Kodama T, Akeda Y, Kono G, Takahashi A, Imura K, et al. (2002) The EspB protein of enterohaemorrhagic *Escherichia coli* interacts directly with  $\alpha$ -catenin. *Cell Microbiol* 4: 213–222.
- Hamaguchi M, Hamada D, Suzuki KN, Sakata I, Yanagihara I. (2008) Molecular basis of actin reorganization promoted by binding of enterohaemorrhagic *Escherichia coli* EspB to  $\alpha$ -catenin. *FEBS J* 275: 6260–6267.
- Knappstein S, Ide T, Schmidt MA, Heussipp G. (2004)  $\alpha$ 1-Antitrypsin binds to and interferes with functionality of EspB from atypical and typical enteropathogenic *Escherichia coli* strains. *Infect Immun* 72: 4344–4350.
- Iizumi Y, Sagara H, Kabe Y, Azuma M, Kume K, et al. (2007) The enteropathogenic *E. coli* effector EspB facilitates microvillus effacing and antiphagocytosis by inhibiting myosin function. *Cell Host Microbe* 2: 383–392.
- Hartland EL, Daniell SJ, Delahay RM, Neves BC, Wallis T, et al. (2000) The type III protein translocation system of enteropathogenic *Escherichia coli* involves EspA-EspB protein interactions. *Mol Microbiol* 35: 1483–1492.
- Ide T, Laarmann S, Greune L, Schillers H, Oberleithner H, et al. (2001) Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic *Escherichia coli*. *Cell Microbiol* 3: 669–679.
- Taylor KA, Luther PW, Donnenberg MS. (1999) Expression of the EspB protein of enteropathogenic *Escherichia coli* within HeLa cells affects stress fibers and cellular morphology. *Infect Immun* 67: 120–125.
- Hamada D, Kato T, Ikegami T, Suzuki KN, Hayashi M, et al. (2005) EspB form enterohaemorrhagic *Escherichia coli* is a natively partially folded protein. *FEBS J* 272: 756–768.
- Kuwajima K. (1996) The molten globule state of  $\alpha$ -lactalbumin. *FASEB J* 10: 102–109.
- Kuwajima K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* 6: 87–103.
- Ohgushi M, Wada A. (1983) ‘Molten-globule state’: a compact form of globular proteins with mobile side-chains. *FEBS Lett* 164: 21–24.
- Uversky VN, Ptitsyn OB. (1996) Further evidence on the equilibrium ‘premolten globule state’: four-state guanidinium chloride-induced unfolding of carbonic anhydrase B at low temperature. *J Mol Biol* 255: 215–228.
- Stryer L. (1965) The interaction of a naphthalene dye with apomyoglobin and apohemoglobin: A fluorescent probe of non-polar binding sites. *J Mol Biol* 13: 482–495.
- Xue B, Williams RW, Oldfield CJ, Goh GK, Dunker AK, et al. (2010) Viral disorder or disordered viruses: do viral proteins possess unique features? *Protein Pept Lett* 17: 932–951.
- Bradford MM. (1976) Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254.
- Sreerama N, Woody RW. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287: 252–260.
- Hamada D, Kuroda Y, Kataoka M, Aimoto S, Yoshimura T et al. (1996) Role of heme axial ligands in the conformational stability of the native and molten globule states of horse cytochrome *c*. *J Mol Biol* 256: 172–186.
- Semenyuk AV, Svergun DI. (1991) GNOM - a program package for small-angle scattering data processing. *J Appl Cryst* 24: 537–540.
- Svergun D, Barberato C, Koch MHJ. (1995) CRYSOLE - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Cryst* 28: 768–773.
- Benning MM, Hong SB, Raushel FM, Holden HM. (2000) The binding of substrate analogs to phosphotriesterase. *J Biol Chem* 275: 30556–30560.

22. Yutani K, Ogasahara K, Kuwajima K. (1992) Absence of the thermal transition in apo- $\alpha$ -lactalbumin in the molten globule state. A study by differential scanning microcalorimetry. *J Mol Biol* 228: 347–350.
23. Geierhaas CD, Nickson AA, Lindorff-Larsen K, Clarke J, Vendruscolo M. (2007) BPPred: A computational tool to predict biophysical quantities of proteins. *Protein Sci* 16: 125–134.
24. Believer D. Biophysical characterization of intrinsically disordered proteins. (2009) *Curr Opin Struct Biol* 19: 23–30.
25. Koch MH, Vachette P, Svergun DI. (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys* 36: 147–227.
26. Doniach S. (2001) Changes in biomolecular conformation seen by small angle X-ray scattering. *Chem Rev* 101: 1763–1778.
27. Kataoka M, Goto Y. (1996) X-ray solution scattering studies of protein folding. *Fold Des* 1: R107–R114.
28. Khon JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, et al. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc Natl Acad Sci USA* 101: 12491–12496.
29. Li X, Romero P, Rani M, Dunker AK, Obradovic Z. (1999) Predicting protein disorder for N-, C- and internal regions. *Genome Inform* 10: 30–40.
30. Romero P, Obradovic Z, Dunker AK. (1997) Sequence data analysis for long disordered regions prediction in the calcineurin family. *Genome Inform* 8: 110–124.
31. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, et al. (2001) Sequence complexity of disordered protein. *Proteins* 42: 38–48.
32. Dosztányi Z, Csizmók V, Tompa P, Simon I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21: 3433–3434.
33. Hirose S, Shimizu K, Kanai S, Kuroda Y, Noguchi T. (2007) POODLE-L: a two-level SVM prediction system for reliably predicting long disordered regions. *Bioinformatics* 23: 2046–2053.
34. Lau SYM, Taneja AK, Hodges RS. (1984) Synthesis of a model protein of defined secondary and quaternary structure. *J Biol Chem* 259: 13253–13261.
35. Muñoz V, Serrano L. (1994) Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct Biol* 1: 399–409.
36. Muñoz V, Serrano L. (1994) Elucidating the folding problem of  $\alpha$ -helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J Mol Biol* 245: 275–296.
37. Muñoz V, Serrano L. (1994) Elucidating the folding problem of  $\alpha$ -helical peptides using empirical parameters III: Temperature and pH dependence. *J Mol Biol* 245: 297–308.
38. Muñoz V, Serrano L. (1997) Development of the multiple sequence approximation within the Agadir model of  $\alpha$ -helix formation. Comparison with Zimm-Bragg and Lifson-Roig formalisms. *Biopolymers* 41: 495–509.
39. Lacroix E, Viguera AR, Serrano L. (1998) Elucidating the folding problem of  $\alpha$ -helices: Local motifs, long-range electrostatics, ionic strength dependence and prediction of NMR parameters. *J Mol Biol* 284: 173–191.
40. Chaffotte AF, Guijarro JI, Guillou Y, Delepiere M, Goldberg ME. (1997) The “pre-molten globule,” a new intermediate in protein folding. *J Protein Chem* 16: 433–439.
41. Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN et al. (2005) Coupling folding and binding with  $\alpha$ -helix-forming molecular recognition elements. *Biochemistry* 44: 12454–12470.
42. Hammers GG, Chang Y-C, Oas TG. (2009) Conformational selection or induced fit: A flux description of reaction mechanism. *Proc Natl Acad Sci USA*, 106: 13737–13741.
43. Kiefhaber T, Bachmann A, Jensen KS. (2012) Dynamics and mechanism of coupled protein folding and binding reactions. *Curr Opin Struct Biol* 22: 21–29.
44. Dyson HJ, Wright PE. (2002) Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* 12: 54–60.
45. Onitsuka M, Kamikubo H, Yamazaki Y, Kataoka M. (2008) Mechanism of induced folding: Both folding before binding and binding before folding can be realized in staphylococcal nuclease mutants. *Proteins* 72: 837–847.
46. Yu Q, Ye W, Wang W, Chen H-F. (2013) Global conformational selection and local induced fit for recognition between intrinsic disordered p53 and CBP. *PLoS ONE* 8: e59627.
47. Chatterjee S, Chaudhury S, McShan AC, Kaur K, De Guzman RN. (2013) Structure and biophysics of type III secretion in bacteria. *Biochemistry* 52: 2508–2517.
48. Wilkins D, Grimshaw SB, Reveur V, Dobson CM, Jones JA et al. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38: 16424–16431.
49. Cole C, Barber JD, Barton GJ. (2008) The Jpred3 secondary structure prediction server. *Nucleic Acids Res* 35: W197–W201.