

# Combining crystallography and EPR: crystal and solution structures of the multidomain cochaperone DnaJ

Thomas R. M. Barends,<sup>a\*</sup>  
Richard W. W. Brosi,<sup>b‡</sup> Andrea  
Steinmetz,<sup>a‡</sup> Anna Scherer,<sup>a</sup>  
Elisabeth Hartmann,<sup>a</sup> Jessica  
Eschenbach,<sup>a</sup> Thorsten Lorenz,<sup>a</sup>  
Ralf Seidel,<sup>c</sup> Robert L. Shoeman,<sup>a</sup>  
Sabine Zimmermann,<sup>a</sup> Robert  
Bittl,<sup>b</sup> Ilme Schlichting<sup>a</sup> and  
Jochen Reinstein<sup>a</sup>

<sup>a</sup>MPI for Medical Research, Heidelberg, Germany, <sup>b</sup>Department of Physics, Freie Universität Berlin, Berlin, Germany, and <sup>c</sup>MPI for Molecular Physiology, Dortmund, Germany

‡ These authors contributed equally.

Correspondence e-mail:  
thomas.barends@mpimf-heidelberg.mpg.de

Hsp70 chaperones assist in a large variety of protein-folding processes in the cell. Crucial for these activities is the regulation of Hsp70 by Hsp40 cochaperones. DnaJ, the bacterial homologue of Hsp40, stimulates ATP hydrolysis by DnaK (Hsp70) and thus mediates capture of substrate protein, but is also known to possess chaperone activity of its own. The first structure of a complete functional dimeric DnaJ was determined and the mobility of its individual domains in solution was investigated. Crystal structures of the complete molecular cochaperone DnaJ from *Thermus thermophilus* comprising the J, GF and C-terminal domains and of the J and GF domains alone showed an ordered GF domain interacting with the J domain. Structure-based EPR spin-labelling studies as well as cross-linking results showed the existence of multiple states of DnaJ in solution with different arrangements of the various domains, which has implications for the function of DnaJ.

Received 14 February 2013  
Accepted 18 April 2013

PDB References: DnaJ, 4j7z;  
4j80

## 1. Introduction

DnaJ protects the cell from adverse conditions through the prevention of protein aggregation or by assisting the refolding of unfolded proteins (Hendrick *et al.*, 1993; Hendrick & Hartl, 1993; Karzai & McMacken, 1996) and is an essential partner of the chaperone DnaK in bacterial DnaK–DnaJ–GrpE chaperone systems. In eukaryotes, corresponding systems exist in which homologues of the DnaJ proteins, the Hsp40 proteins, function together with Hsp70 proteins, which are the eukaryotic variant of DnaK (Caplan *et al.*, 1993; Cyr *et al.*, 1994; Rassow *et al.*, 1995; Aron *et al.*, 2005; Qiu *et al.*, 2006). While sequence homologues of the nucleotide-exchange factor GrpE are missing in eukaryotes, their role is taken over by functional equivalents such as Hop.

In the DnaK ATPase cycle, DnaJ stimulates ATPase activity, causing DnaK to switch from the ATP-bound state which binds substrates weakly to an ADP-bound state which binds substrates tightly. GrpE then stimulates the exchange of ADP for ATP, resetting the cycle. Thus, in performing these ATPase-related functions, DnaJ and GrpE are also deeply involved in regulating substrate binding by DnaK. To this end, their respective activities are finely balanced to achieve optimal chaperone activity (Pierpaoli *et al.*, 1998).

In contrast to DnaJ from *Escherichia coli* or *Methanobacter thermotrophicus*  $\Delta$ H (Laufen *et al.*, 1999; Popp & Reinstein, 2009; Russell *et al.*, 1999), DnaJ from *Thermus thermophilus* stimulates the ATPase activity of DnaK only very weakly (Groemping *et al.*, 2001; Klostermeier

*et al.*, 1999; Motohashi *et al.*, 1997). Nevertheless, DnaJ is essential and highly effective in preventing aggregation of challenged protein-folding intermediates as part of the *T. thermophilus* DnaK–DnaJ–GrpE chaperone system.

The prevention of protein aggregation would appear to be especially important and difficult to achieve in organisms living at very high temperatures, such as *T. thermophilus*. Consequently, the *T. thermophilus* DnaK–DnaJ chaperone system has attracted much interest (Motohashi *et al.*, 1994, 1996; Klostermeier *et al.*, 1999; Schlee & Reinstein, 2002). It has been shown that *T. thermophilus* DnaK and *T. thermophilus* DnaJ combine with the assembly factor DafA to form a DnaK<sub>3</sub>DnaJ<sub>3</sub>DafA<sub>3</sub> complex, from which DafA is expelled by the substrate protein, which binds to DnaK (Motohashi *et al.*, 1996; Dumitru *et al.*, 2004; Watanabe & Yoshida, 2004). In the presence of DafA, *T. thermophilus* DnaJ reduces substrate binding by DnaK and even replaces substrate bound by DnaK (Klostermeier *et al.*, 1999).

However, the precise function and mode of action of *T. thermophilus* DnaJ, or of DnaJ in general, is unclear. Structural information could help to remedy this, but such data have been unavailable to date as neither the structure of any full-length DnaJ nor of a domain of *T. thermophilus* DnaJ has been determined. However, structures and sequences of domains of related proteins help to formulate some hypotheses. Like every DnaJ, *T. thermophilus* DnaJ is a multidomain protein which possesses the highly conserved J domain with its characteristic HPD motif (Bork *et al.*, 1992; Wall *et al.*, 1994; Szyperski *et al.*, 1994). *T. thermophilus* DnaJ is a typical type II DnaJ (see Walsh *et al.*, 2004 for a definition of DnaJ subfamilies), *i.e.* it lacks the zinc-finger domains found in type I DnaJs such as that of *E. coli*. Importantly, however, *T. thermophilus* DnaJ can substitute for *E. coli* DnaJ in *in vitro* refolding experiments with the *E. coli* DnaK–DnaJ system, showing that *T. thermophilus* DnaJ is a full representative of the DnaJ family and that zinc-finger domains are not required for cochaperone activity (Groemping *et al.*, 2005).

In both type II and type I DnaJs the highly conserved J domain is found at the N-terminus. Biochemical studies on the *E. coli* DnaK–DnaJ system have shown that in this system the J domain, and in particular its HPD motif, is important for both DnaK binding and ATPase stimulation (Wall *et al.*, 1994; Karzai & McMacken, 1996). Thus, by analogy, in *T. thermophilus* DnaJ the J domain may also serve to interact with DnaK. Directly behind the J domain, *T. thermophilus* DnaJ possesses a polyproline motif consisting of six consecutive proline residues with unknown function. The polyproline motif is followed by a glycine/phenylalanine-rich domain or GF domain as is often observed in DnaJ proteins (Wall *et al.*, 1995). Because of their high glycine content, the GF domains of DnaJ proteins have sometimes been suggested to serve as a flexible linker between the J domain and the C-terminal domain. However, Karzai & McMacken (1996) found that both the J and the GF domains are required for DnaK ATPase activation in the *E. coli* system, whereas the J domain on its own was incapable of doing so. Also in *E. coli*, the GF domain was found to be essential for the activation of the ATP-

mediated substrate binding of DnaK (Wall *et al.*, 1995). Moreover, in a eukaryotic system, yeast, the GF domain determines the binding specificity of the Hsp40 Sis1 (Yan & Craig, 1999). The GF domain is in turn followed by a C-terminal domain of variable sequence.

The first structural information on a DnaJ domain was obtained from NMR studies of the *E. coli* DnaJ J domain alone and together with the GF domain (Szyperski *et al.*, 1994; Pellecchia *et al.*, 1996). While the GF domain was too disordered for an atomic model to be built, the J domain was shown to consist of four helices labelled I–IV, where the N-terminal helix I is connected *via* a short loop to helices II and III, forming a helix–turn–helix motif with the latter two helices in a relative orientation typical of a coiled-coil structure. The HPD motif is positioned in the loop between helices II and III, which was found to be poorly ordered. Helix III is connected to helix IV by a short linker making an  $\sim 90^\circ$  turn between the two helices. Helix IV is followed by a poorly ordered linker to the GF domain. Studies of a construct with both the J and GF domains showed that although the GF domain was also disordered, the core of this domain is not completely disordered in the *E. coli* protein as indicated by relaxation times. This partial order within the GF domain was proposed to offset the entropic disadvantage of participating in complex formation with DnaK when starting with a completely disordered state with very high entropy (Pellecchia *et al.*, 1996). Another NMR study of *E. coli* DnaJ domains showed that upon including the GF domain in a construct with the J domain, the dynamics of the J domain change measurably. Including the GF domain in the construct particularly changed the dynamics of the HPD motif aspartate (Huang, Ghose *et al.*, 1999). The same laboratory also reported that inclusion of the GF domain changes the orientation of helix IV and suggested that the GF domain interacts transiently with the J domain, proposing that this interaction takes place with the surface-exposed area of helix III opposite its interface with helix II (Huang, Flanagan *et al.*, 1999).

Despite their low sequence identity, the few available structures of C-terminal domains of DnaJs and Hsp40s are remarkably similar (Sha *et al.*, 2000; Li *et al.*, 2006; Hu *et al.*, 2008; Midwest Center for Structural Genomics, unpublished work; Suzuki *et al.*, 2010), consisting of two  $\beta$ -barrel-like subdomains stacked on top of each other into a rod-like structure 100 Å in length. Two C-terminal domains dimerize by domain swapping involving two helices, resulting in a laundry-pin-like structure, as can be seen in crystal structures of the C-terminal domains of the yeast Hsp40 Sis1 (Li *et al.*, 2006). Recently, the structure of the monomeric endoplasmic reticulum type III DnaJ protein P58IPK has been reported (Svård *et al.*, 2011), which has the J domain at the C-terminus, following an  $\alpha$ -helical N-terminal stalk structurally unrelated to the dimerization domains mentioned above.

Here, we describe an in-depth structural characterization of *T. thermophilus* DnaJ by hybrid methods. Crystals of wild-type *T. thermophilus* DnaJ were grown and their diffraction was improved to 3.8 Å resolution using a dehydration protocol, but they remained resistant to phasing attempts. A divide-and-

conquer strategy led us to first determine the structure of a construct containing the N-terminal 114 residues, DnaJTth<sub>114</sub>, comprising the J and GF domains, using radiation-damage-induced phasing with anomalous scattering (RIPAS; Ravelli *et al.*, 2003; Zwart *et al.*, 2004; Banumathi *et al.*, 2004; Ramagopal *et al.*, 2005). The electron-density map of this structure inspired a third construct, this time of a complete DnaJ from which a disordered stretch of seven amino acids (108–114 in the linker between the GF domain and C-terminal domain; DnaJ<sub>Δ108–114</sub>) was deleted with a view to increasing conformational stability. The subsequent substitution of four residues with methionine (construct DnaJ<sub>Δ108–114/4M</sub>) allowed SeMet-SAD phasing, resulting in the determination of the first crystal structure of a complete functional DnaJ containing J, GF and C-terminal domains to 2.9 Å resolution. This structure was then used as a search model for molecular replacement with data from the wild-type DnaJ crystals, which indicated a large degree of flexibility between the GF and C-terminal domains, which was then investigated by EPR spin-labelling, simple modelling and cross-linking experiments, resulting in a hybrid model of the highly dynamic DnaJ molecule.

## 2. Materials and methods

### 2.1. Protein expression and purification

Constructs were prepared by either QuikChange or two-stage PCR. After two-stage PCR, the PCR products were ligated into pCR2.1TOPO (Invitrogen) and positive clones were selected using blue/white screening. Correct clones as identified by sequencing were digested with *NcoI* and *NotI* (New England Biolabs), purified by agarose-gel electrophoresis and ligated into pET24d. For the DnaJ<sub>Δ108–114/4M</sub> construct, the region coding for residues 108–114 was deleted and residues Leu57, Ile149, Leu180 and Leu233 were mutated to methionine because of their aliphatic nature.

All DnaJ constructs were heterologously expressed in *E. coli* BL21(DE3) cells and were purified essentially as described in Klostermeier *et al.* (1999) either by ammonium sulfate precipitation and ion-exchange chromatography (DnaJTth<sub>114</sub>) or by a combination of heat denaturation and Ni<sup>2+</sup>-affinity and gel-filtration chromatography (other constructs). DnaJTth<sub>114</sub> was expressed in *E. coli* BL21(DE3) cells and precipitated from a cell-free extract by adding solid ammonium sulfate to 80% saturation at room temperature, followed by stirring for 30 min at 277 K. The precipitate was recovered by ultracentrifugation (30 min at 277 K at 65 000g) and dissolved in buffer *B* (50 mM Tris-HCl, 25 mM KCl, 10% glycerol, 2 mM EDTA, 1 mM DTT pH 7.5 at 298 K). Further contaminants were heat-denatured in a water bath at 348 K for 30 min, followed by removal of the precipitated material using ultracentrifugation (30 min, 65 000g, 277 K). The supernatant was loaded onto a Q Sepharose column equilibrated with buffer *B*. DnaJTth<sub>114</sub> was found in the flowthrough and its purity was assessed by SDS-PAGE. When required, the protein was further purified on a Superdex 75 column in buffer *C* (50 mM Tris-HCl, 200 mM KCl, 10% glycerol, 2 mM EDTA, 1 mM

DTT pH 7.5 at 298 K). The mass of the protein as determined by MALDI-TOF MS was 12 902.7 Da, which is nearly identical to the calculated average molecular mass of the construct minus the N-terminal methionine residue (12 903.4 Da).

For the production of most of the other constructs, *E. coli* BL21(DE3) cells were transformed with the appropriate plasmid and grown on 3–6 l LB medium. For the production of SeMet-DnaJ<sub>Δ108–114/4M</sub>, 5 l minimal medium supplemented with 50 mg l<sup>-1</sup> L-selenomethionine at 310 K was used (Doublé, 1997). At an OD<sub>600</sub> of 0.3–0.8, 1 mM IPTG was added and the temperature was lowered to 293 K. After 20 h, the cells were harvested by centrifugation for 10 min at 5000 rev min<sup>-1</sup> in an SLC-6000 rotor at 277 K. The cells were resuspended in 50 ml buffer *D* (50 mM sodium phosphate, 10 mM imidazole, 300 mM sodium chloride pH 8.0) and a Complete EDTA-free protease-inhibitor cocktail tablet was added. The cells were lysed by two passes through a fluidizer (690 kPa), and the lysate was incubated in a water bath at 343 K for 10 min and then cooled on ice. The mixture was centrifuged for 40 min at 30 000 rev min<sup>-1</sup> in a Beckman Ti60 ultracentrifuge rotor (277 K) and the supernatant was mixed with 12 ml 50% Ni-NTA agarose in buffer *D*. The resulting suspension was gently agitated for 2 h at 277 K and then poured into a 6 ml column. This column was washed with 60 ml buffer *E* (50 mM sodium phosphate, 20 mM imidazole, 300 mM sodium chloride pH 8.0). The protein was then eluted with 12 ml buffer *F* (50 mM sodium phosphate, 500 mM imidazole, 300 mM sodium chloride pH 8.0). The protein was concentrated by ultrafiltration to 3 ml and submitted to a 340 ml Superdex 200 column equilibrated with buffer *G* (25 mM bis-tris, 200 mM KCl, 10% glycerol pH 6.5 adjusted with HCl) running at 2 ml min<sup>-1</sup>. Pure fractions as judged by SDS-PAGE analysis were pooled and then washed and concentrated by ultrafiltration in buffer *H* (20 mM bis-tris, 25 mM KCl pH 6.5 adjusted with HCl) to an A<sub>280</sub><sup>1 cm</sup> of ~9–14. The protein was flash-cooled in liquid nitrogen and stored at 193 K.

### 2.2. Crystallization and structure determination

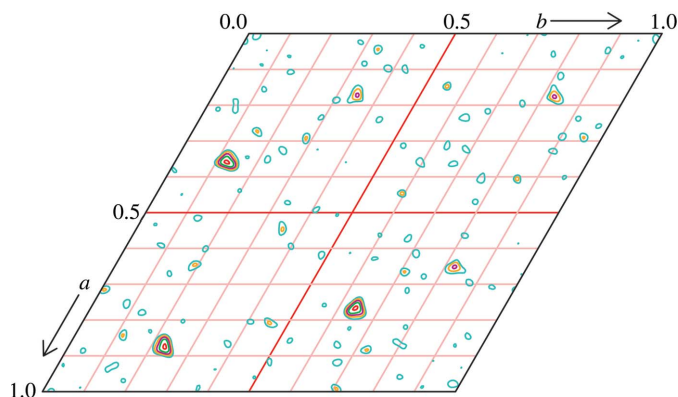
**2.2.1. J/GF-domain construct DnaJTth<sub>114</sub>.** DnaJTth<sub>114</sub> was crystallized by equilibrating hanging drops consisting of 2 µl 17 mg ml<sup>-1</sup> DnaJTth<sub>114</sub> in buffer *B* plus 10 mM SrCl<sub>2</sub> and 2 µl 35% PEG 1500 against a 600 µl reservoir of 35% PEG 1500 in Linbro plates. Crystals were cryoprotected in 35% PEG 1500, 10% PEG 400, 25 mM Tris-HCl pH 7.5, 10 mM SrCl<sub>2</sub>. All data were processed with *XDS* (Kabsch, 1993) or *DENZO/SCALEPACK* (Otwinowski & Minor, 1997; Otwinowski, 1993). A mercury derivative was prepared by soaking a crystal in cryoprotectant solution for 16 h which contained 10 mM HgCl<sub>2</sub> instead of SrCl<sub>2</sub>, followed by back-soaking for 1 min. A 360° data set to 2.5 Å resolution was collected on a Rigaku MicroMax-007 HF rotating-anode generator equipped with Osmic mirrors and a MAR345 image plate using Cu K $\alpha$  radiation. These data sets showed appreciable anomalous signal, but were nonisomorphous with several other data sets, and SAD phasing was unsuccessful. Since mercury derivatives

**Table 1**  
Data-collection, phasing and refinement statistics.

	DnaJ <sub>114</sub> , 0–120° data set, HgCl <sub>2</sub> derivative	DnaJ <sub>114</sub> , 240–360° data set, HgCl <sub>2</sub> derivative (pseudo-native)	DnaJ <sub>114</sub> , high-resolution data set (4j7z)	DnaJ <sub>Δ108–114/4M</sub> , SAD data set	DnaJ <sub>Δ108–114/4M</sub> , high-resolution data set (4j80)	DnaJ WT, low resolution
Data collection						
Space group	<i>P</i> 3 <sub>1</sub>		<i>P</i> 3 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 84.2, <i>c</i> = 71.9		<i>a</i> = <i>b</i> = 86.0, <i>c</i> = 72.7	<i>a</i> = 102.8, <i>b</i> = 105.3, <i>c</i> = 130.6	<i>a</i> = 102.1, <i>b</i> = 105.0, <i>c</i> = 130.0	<i>a</i> = 75.1, <i>b</i> = 131.6, <i>c</i> = 153.1
Resolution (Å)	40–2.50 (2.60–2.50)		20–1.64 (1.70–1.64)	80–3.20 (3.30–3.20)	20–2.90 (3.00–2.90)	40–3.80 (3.94–3.80)
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub> ( <i>I</i> /σ( <i>I</i> ))	0.044 (0.197)†	0.041 (0.194)†	0.061 (0.317)	0.086 (0.370)†	0.063 (0.415)	0.104 (0.564)
Completeness (%)	87.4 (73.4)	85.9 (72.2)	94.5 (82.5)	100 (100)	99.2 (99.8)	99.1 (97.9)
Multiplicity	1.9 (1.9)	2.0 (1.9)	3.1 (2.5)	7.8 (7.8)	4.1 (4.1)	8.0 (7.5)
Refinement						
Resolution (Å)			20–1.64		20–2.90	
No. of reflections			68483		29725	
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>			0.214/0.246		0.291/0.328	
No. of atoms						
Protein			4921‡		8460§	
Ligand/ion			18 [glycerol]		0	
Water			341		0	
<i>B</i> factors (Å <sup>2</sup> )						
Protein			35		25	
Ligand/ion			22 [glycerol]			
Water			33			
R.m.s. deviations						
Bond lengths (Å)			0.010		0.010	
Bond angles (°)			1.17		1.15	

† Considering Friedel mates as individual reflections. ‡ Six monomers in the asymmetric unit: chains *A*, *B* and *F*, residues 2–101; chain *C*, residues 2–104; chain *D*, residues 2–107; chain *E*, residues 4–101. § Two dimers in the asymmetric unit: chains *A*, *B*, *C* and *D* containing residues 2–272.

are prone to radiation damage (Ramagopal *et al.*, 2005), we set out to phase the data using radiation-damage-induced phasing with anomalous scattering (RIPAS; Ravelli *et al.*, 2003; Zwart *et al.*, 2004; Banumathi *et al.*, 2004; Ramagopal *et al.*, 2005). The data were reprocessed in two parts (see Table 1). The first 120° of data were expected to have suffered the least from radiation damage and were used as a derivative. The last 120°, *i.e.* from  $\varphi = 240^\circ$  to  $\varphi = 360^\circ$ , were expected to have suffered the most from radiation damage and were used as a pseudo-native. Using these data sets, an isomorphous difference



**Figure 1**  
Super-sharpened origin-removed isomorphous difference Patterson section ( $w = 2/3$ ) calculated between the final 120° and the first 120° of data of a 360° data set of the mercury derivative, showing the effect of radiation damage on the mercury derivative. Contour interval =  $1\sigma$ . This figure was prepared with *XPREP* (Schneider & Sheldrick, 2002).

Patterson calculated with *XPREP* (Schneider & Sheldrick, 2002; Fig. 1) showed peaks in the same positions as in the anomalous difference Patterson map for the whole 360° data set. *SHARP* (Vonrhein *et al.*, 2007) was used to obtain phases by combining the isomorphous differences between the two data sets with the anomalous difference information, which resulted in an interpretable map (FOM of 0.20 before solvent flattening for all reflections). Using *Xfit* (McRee, 1999), six copies of the J domain were identified in the electron density. Sixfold averaging and solvent flattening using *RESOLVE* (Terwilliger, 2003, 2004) resulted in a much improved map (FOM of 0.61 for all reflections) into which most of the sequence could be built. The final model was obtained through iterative cycles of model building in *Xfit* (McRee, 1999) and either simulated annealing with *CNS* (Brünger *et al.*, 1998) or refinement with *REFMAC5* (Murshudov *et al.*, 2011), using riding H atoms and TLS refinement, against a high-resolution data set measured on the SLS X10SA beamline at 90 K and a wavelength of 0.9536 Å. Because of the sixfold noncrystallographic symmetry, the reflections used for the calculation of *R*<sub>free</sub> were selected in thin resolution shells to avoid contamination of the test set. The geometry of the final model was excellent as witnessed by the Ramachandran plot statistics, with 93.9% of the residues in the core regions, 5.9% and 0.2% in the allowed and generously allowed regions, respectively, and 0.0% in disallowed regions. Data-collection and model statistics are given in Table 1.

**2.2.2. DnaJ<sub>Δ108–114/4M</sub>.** Crystals of SeMet-DnaJ<sub>Δ108–114/4M</sub> were grown by the hanging-drop vapour-diffusion method,

equilibrating a drop consisting of 2  $\mu\text{l}$  protein solution with 2 mM of various adenine nucleotides and 2  $\mu\text{l}$  reservoir solution against a reservoir consisting of 14–16% PEG 6000, 0.1 M succinate/BTP buffer pH 7.0. Crystals were flash-cooled in liquid nitrogen after cryoprotection in 25% ethylene glycol, 16% PEG 6000, 0.1 M succinate/BTP buffer pH 7.0 and stored in liquid nitrogen.

A 3.2 Å resolution SAD data set was collected from a crystal grown in the presence of 2 mM AMPPNP on the SLS X10SA beamline at a temperature of 90 K and a wavelength of 0.97893 Å. On the same beamline, a 2.9 Å resolution data set was collected at 90 K and a wavelength of 1.00767 Å, *i.e.* at an energy below the Se absorption edge, from a crystal grown in the presence of 2 mM ATP-NH<sub>2</sub>. Selenium positions were found using *SHELXD* (Sheldrick, 2008). Phases were calculated using *autoSHARP* (Vonrhein *et al.*, 2007), resulting in an interpretable map (phasing power 1.67, FOM<sub>acentric</sub> = 0.40, FOM<sub>centric</sub> = 0.13 for the entire resolution range). Inclusion of the 2.9 Å resolution data and solvent flattening with *DM* resulted in an improved map into which the structure was built by iterative cycles of rebuilding using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010) and refinement with *CNS* (Brünger *et al.*, 1998) and *REFMAC5* (Murshudov *et al.*, 2011). Noncrystallographic symmetry restraints and TLS refinement were used. The anomalous signal present in the SAD data set was used to calculate an anomalous difference Fourier map which was used to locate the Se atoms to guide model building. For the J domains of chains *B* and *D* the  $2mF_o - DF_c$  density (Read, 1986) was poor but recognizable such that a J domain could be placed there. A peak in the anomalous difference electron-density map for the lone Se atom of the labelled J domains confirmed their locations. The final model displayed good geometry, with 91.6% of the residues in the core regions of the Ramachandran plot, 7.8% and 0.6% in the allowed and generously allowed regions, respectively, and 0.0% in disallowed regions. Crystal structures and diffraction data have been deposited in the PDB with codes 4j7z (DnaJTh<sub>114</sub>) and 4j80 (DnaJ<sub>Δ108–114/4M</sub>).

**2.2.3. Wild-type DnaJ.** Wild-type (WT) DnaJ crystals were produced using the sitting-drop method by equilibrating 2  $\mu\text{l}$  protein solution in buffer *H* with an  $A_{280}^{1\text{cm}}$  of ~9–14 plus 2  $\mu\text{l}$  reservoir solution and 0.2  $\mu\text{l}$  24 mM CYMAL-5 (Hampton Research, Aliso Viejo, USA) against 600  $\mu\text{l}$  reservoir solution consisting of 1.0 M (NH<sub>4</sub>)<sub>2</sub>HPO<sub>4</sub>, 0.1 M Tris-HCl pH 8.5. The resulting box-shaped crystals were dehydrated overnight in a saturated solution of trimethylamine oxide (Sigma) in reservoir solution, which served as both a dehydrating agent and a cryoprotectant. Typically, the diffraction resolution improved to around 4 Å after 4 h of dehydration. Dehydration for longer than 20 h reduced the resolution. The best data (3.8 Å resolution) were obtained at the SLS from a crystal which had been soaked for 16 h in 3 mM *p*-chloromercuribenzoate in the dehydration solution. The resulting data were phased by molecular replacement with *Phaser* (McCoy *et al.*, 2005) using a monomer of the C-terminal domain of the SeMet-DnaJ<sub>Δ108–114/4M</sub> structure as a search model. Only C-terminal domains were observed in the density, and neither molecular

replacement nor phased molecular replacement with models constructed from the DnaJTh<sub>114</sub> structure resulted in localization of the J domains. No models were refined.

### 2.3. Spin labelling and EPR measurements

For spin labelling of DnaJ variants (Todd *et al.*, 1987, 1989), the proteins were purified as usual but without a heat-precipitation step in order to minimize cysteine oxidation. Immediately after elution from the Ni-NTA column, the protein was rapidly mixed with a tenfold molar excess of the spin label MTSL [*S*-(2,2,5,5-tetramethyl-2,5-dihydro-1*H*-pyrrol-3-yl)methyl methanesulfonylthioate; Pannier *et al.*, 2000; Jeschke & Polyhach, 2007] by rapidly pipetting the protein into a 15 ml tube containing a fivefold molar excess of MTSL dissolved in ~50  $\mu\text{l}$  acetonitrile while rapidly vortexing, followed by incubation overnight at 277 K. An NAP-25 column (GE Healthcare) was used to remove most of the excess label, after which the purification was continued as usual. After purification, successful labelling was confirmed using MALDI-TOF mass spectrometry. When required, protein samples were deuterated in heavy water using gel filtration and/or ultrafiltration.

Four-pulse ELDOR experiments (also called DEER; double electron–electron resonance; Martin *et al.*, 1998) were measured at X-band frequencies (around 9.4 GHz) on Bruker Elexsys E580 and E680 spectrometers using an MD5 dielectric ring resonator, Oxford Instruments CF935 cryostats and ITC503 temperature controllers. Experiments on samples in protic buffers were conducted at 80 K. Samples in deuterated buffer were measured at 50 K (single mutants), which allows measurements of longer distances owing to prolonged relaxation times (Jeschke & Polyhach, 2007). The ELDOR sample DnaJ monomer concentration was between 200 and 280  $\mu\text{M}$ . The pump and detection pulses were set to 32 and 40 ns with a two-step phase cycle for the initial  $\pi/2$  pulse. The pump pulse was set on the absorption maximum of the nitroxide spectrum and its power was adjusted for maximum inversion efficiency in a two-pulse Hahn echo. Proton modulation was suppressed by an eight-step increase by 8 ns of the initial inter-pulse delay starting at 200 ns. Owing to strong nuclear modulation, this delay was set to 360 ns for the samples in deuterated buffer. The acquired time traces were analyzed using the *MatLab* program *DeerAnalysis* (Jeschke *et al.*, 2006) employing Tikhonov regularization (Tikhonov & Arsenin, 1977) after separating the ELDOR signal from background decay owing to isotropic interaction with other surrounding DnaJ molecules at larger distances. Comparison with crystal structures was performed by calculating spin-label rotamer distributions with the program *MMM* (Polyhach *et al.*, 2011).

### 2.4. Solution cross-linking

The proteins (5 ml each at concentrations of 170–190  $\mu\text{M}$  as determined by UV absorption at 280 nm using  $\epsilon = 23\,505\text{ l mol}^{-1}\text{ cm}^{-1}$ ) were dialysed for 16 h against 1 l buffer *I* (25 mM HEPES, 25 mM KCl, 1 mM EDTA pH 7.5 adjusted

with NaOH). For cross-linking, fresh proteins were diluted to 10  $\mu\text{M}$  in 30 ml buffer *J* (50 mM bis-tris, 10 mM EDTA, 100 mM NaCl pH 6.5 adjusted with NaOH). To these solutions, an equimolar amount of BMB (bismaleimidobutane;

Pierce, Rockford, USA), *i.e.* 1 mol DnaJ dimer:1 mol BMB, was added as follows: 30  $\mu\text{l}$  of 10 mM BMB in DMSO was added in ten steps of 3  $\mu\text{l}$  with vigorous stirring at 5 min intervals. After the last addition the solutions were left to incubate for a further 1.5 h while stirring.

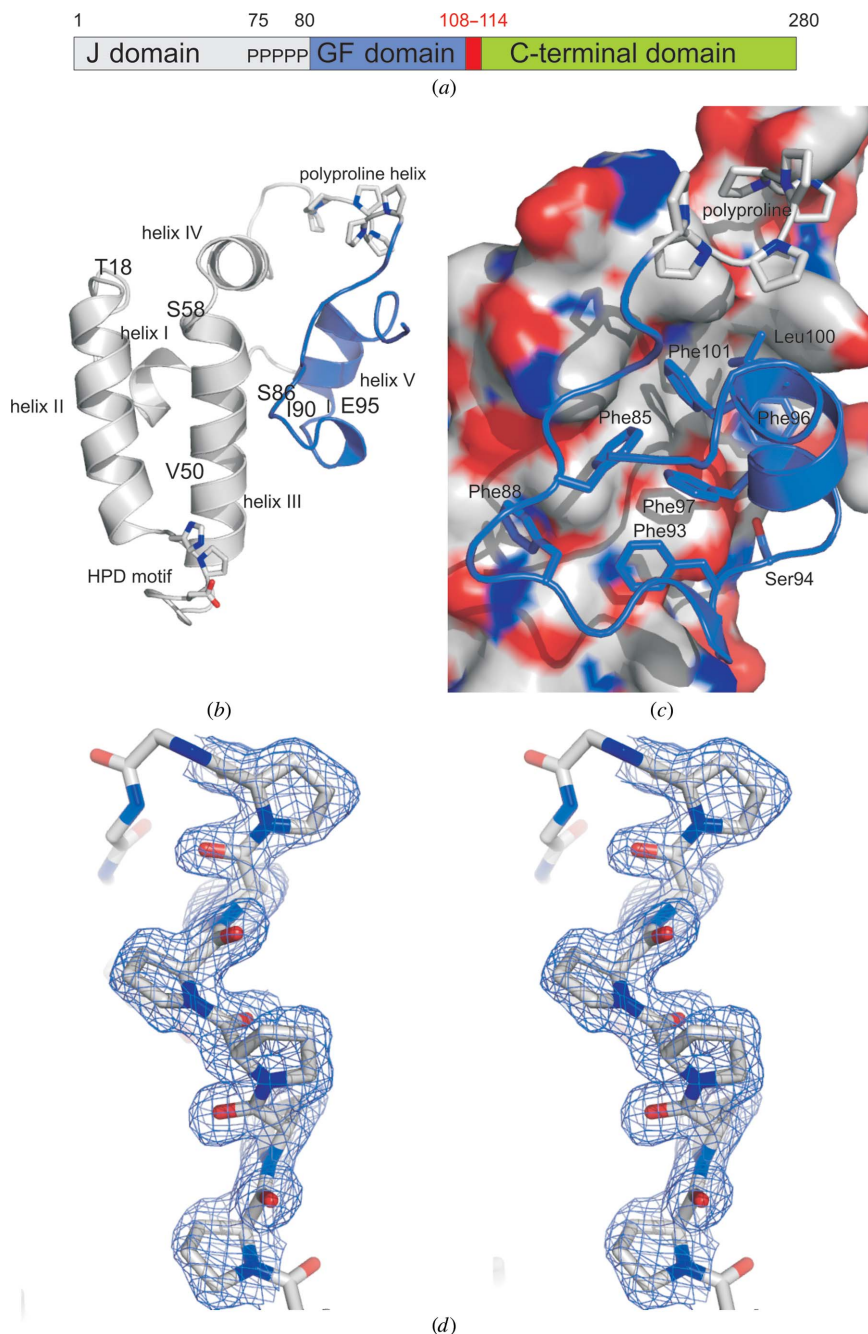
To separate DnaJ dimers from DnaJ tetramers generated by interdimer cross-linking, the proteins were concentrated to 100  $\mu\text{l}$  and submitted to a 23 ml Superdex 12 (10/300) column (GE Healthcare) running at 0.8 ml  $\text{min}^{-1}$  in buffer *K* (50 mM HEPES, 200 mM NaCl, 2 mM DTE). From each mutant, a fraction at the end of the dimer peak (*i.e.* free from tetramer contamination) was investigated using reducing SDS-PAGE to check for the presence of covalently cross-linked DnaJ dimers.

## 2.5. Computational modelling of DnaJ mobility

Computational modelling of the mobility of DnaJ molecules was performed in *MatLab* using simple models of spherical amino acids. Starting from the crystal structure, an ensemble of conformations was calculated using a Monte Carlo algorithm employing a pseudo-energy term calculated from the EPR distance distributions and the number of clashes between residues. Details are included in Appendix C.

## 2.6. Luciferase-refolding activity measurements

DnaK luciferase-refolding assays were performed essentially as described in Beinker *et al.* (2002) and Aponte *et al.* (2010). Briefly, 10  $\mu\text{M}$  luciferase was denatured by incubation for 30 min at room temperature in 25 mM HEPES pH 7.5, 50 mM KCl, 15 mM  $\text{MgCl}_2$ , 10 mM DTE, 0.05 mg  $\text{ml}^{-1}$  BSA, 1 mM ATP with 7 M urea. The denatured luciferase was then diluted 125-fold into 25 mM HEPES pH 7.5, 50 mM KCl, 15 mM  $\text{MgCl}_2$ , 2 mM DTE, 0.05 mg  $\text{ml}^{-1}$  BSA, 1 mM ATP, 10 mM phosphoenolpyruvate, 240  $\mu\text{M}$  coenzyme A, 0.1 mM luciferin, 50  $\mu\text{g ml}^{-1}$  pyruvate kinase, to which 3.2  $\mu\text{M}$  DnaK, 0.4  $\mu\text{M}$  GrpE and 0.16, 0.8 or 3.2  $\mu\text{M}$  DnaJ were added. The refolding of luciferase was then followed for 500 min in microtitre plates using the luminescence at 298 K. Details are given in Appendix A.



**Figure 2**

(a) DnaJ sequence, showing the J (grey), GF (blue) and C-terminal (green) domains, as well as the polyproline region and the deletion in the DnaJ $_{\Delta 108-114}$  mutant (red). (b) Structure of the isolated J/GF domain. The HPD motif and the polyproline region are shown as sticks. EPR spin-labelling and cross-linking sites (Thr18, Ser58, Ser86 and Glu95) are indicated. (c) Interface between the J and GF domains, showing the J-domain surface (coloured by atom type: C, grey; N, blue; O, red) and the GF-domain phenylalanines. All figures use the colour scheme in (a). (d) Stereo image showing the final refined  $2mF_o - DF_c$  electron-density map of the 1.64 Å resolution DnaJ $_{\text{Thr114}}$  structure at  $1\sigma$ . As an example, the polyproline region (residues 75–80) of chain *B* is shown. All molecular-structure figures were prepared using *PyMOL* (DeLano, 2002)

### 3. Results

#### 3.1. The *T. thermophilus* DnaJ J/GF-domain crystal structure shows an ordered GF domain

Based on limited proteolysis of a DnaK–DnaJ–DafA complex (Motohashi *et al.*, 1996; Klostermeier *et al.*, 1999), we first designed a construct containing the first 114 amino acids (DnaJ<sub>114</sub>), comprising the J and GF domains, and determined its structure to 1.64 Å resolution (Figs. 1 and 2) by radiation-damage-induced phasing (Ramagopal *et al.*, 2005; Ravelli *et al.*, 2003; Zwart *et al.*, 2004) from a mercury(II) chloride derivative. Contrary to expectation, the crystal structure showed an ordered GF domain in the same conformation in all six molecules of the asymmetric unit (Figs. 2*b* and 2*c*), which are highly similar (the r.m.s. positional differences between C<sup>α</sup> atoms of any two polypeptide chains range from 0.50 to 0.86 Å for up to 100 C<sup>α</sup> atoms). After helix IV, the polypeptide chain folds back onto the protein at the Pro<sub>6</sub> motif, which adopts a polyproline II conformation (Fig. 2*d*; Adzhubei & Sternberg, 1993). After the Pro<sub>6</sub> motif, the GF domain folds onto one side of the J domain in a spiral, making extensive hydrophobic interactions, mainly *via* the conserved phenylalanine residues. Only two hydrogen bonds are formed between the J and the GF domains: from the Tyr<sub>69</sub> hydroxyl group to the side chain of Glu<sub>99</sub> and between the side chains of Glu<sub>52</sub> and Ser<sub>94</sub>. In correspondence with the NMR observations for the partially disordered *E. coli* DnaJ GF domain (Pellecchia *et al.*, 1996), the *T. thermophilus* DnaJ GF domain binds to helix III of the J domain opposite helix II.

#### 3.2. The *T. thermophilus* DnaJ<sub>Δ108–114</sub> variant is biologically active and yields high-quality crystals

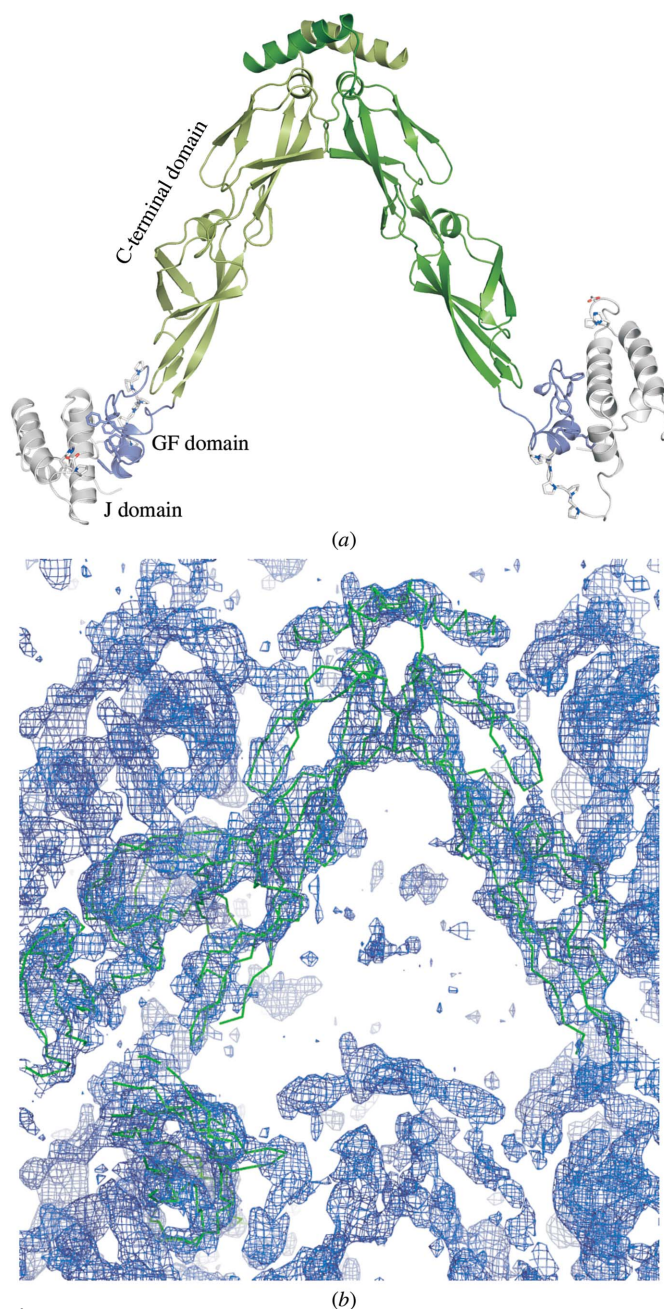
In the DnaJ<sub>114</sub> structure, no electron density was observed for residues 108–114 connecting the GF and the C-terminal domains in the complete protein. Therefore, a full-length construct was prepared that lacked the 108–114 stretch (DnaJ<sub>Δ108–114</sub>), which was shown to be functional in refolding assays (see Fig. 6). Crystallization of this variant allowed structure solution to 2.9 Å resolution using SeMet phasing with a construct containing four additional methionines (DnaJ<sub>Δ108–114/4M</sub>), resulting in a structure containing all three domains (Fig. 3*a*).

#### 3.3. The *T. thermophilus* DnaJ<sub>Δ108–114</sub> crystal structure also displays an ordered GF domain

DnaJ<sub>Δ108–114</sub> dimerizes *via* its C-terminal domains, resulting in a V-shaped molecule with the J/GF domains at the ends of the two stalks (Fig. 1*b*) like feet on a pair of legs. The J and GF domains interact tightly through a hydrophobic interface involving six of the seven phenylalanines of the GF domain and helix III of the J domain (Fig. 1*c*). The same interaction is seen in all six independent copies of the J/GF domains in the DnaJ<sub>114</sub> structure. In all four monomers in the asymmetric unit of the DnaJ<sub>Δ108–114</sub> crystals the orientation of the C-terminal and J/GF domains differs, showing that the Δ108–114 deletion does not restrict their mutual motion. Moreover, the J/GF

regions of chains *B* and *D* showed poor electron density, also indicating flexibility.

We also obtained crystals of wild-type DnaJ which diffracted to 3.8 Å resolution after applying a dehydration protocol. Interestingly, in an electron-density map of wild-type DnaJ phased by molecular replacement with DnaJ<sub>Δ108–114</sub> only the C-terminal domain was observed (Fig. 3*b*), indicating a high degree of mobility of the J and GF domains. Indeed, the packing of the wild-type DnaJ crystals leaves sufficient space



**Figure 3**  
(*a*) Structure of a dimer of DnaJ<sub>Δ108–114</sub> from the 2.9 Å resolution crystal structure. The relative orientations of the J/GF domains and the C-terminal domains differ in the four monomers in the asymmetric unit. (*b*)  $2mF_o - DF_c$  density at  $1.5\sigma$  of the 3.8 Å resolution WT data after molecular replacement with the C-terminal domain of DnaJ<sub>Δ108–114</sub>. No density for the N-terminal domains was found and sufficient space for these domains is present in the packing of the molecules.

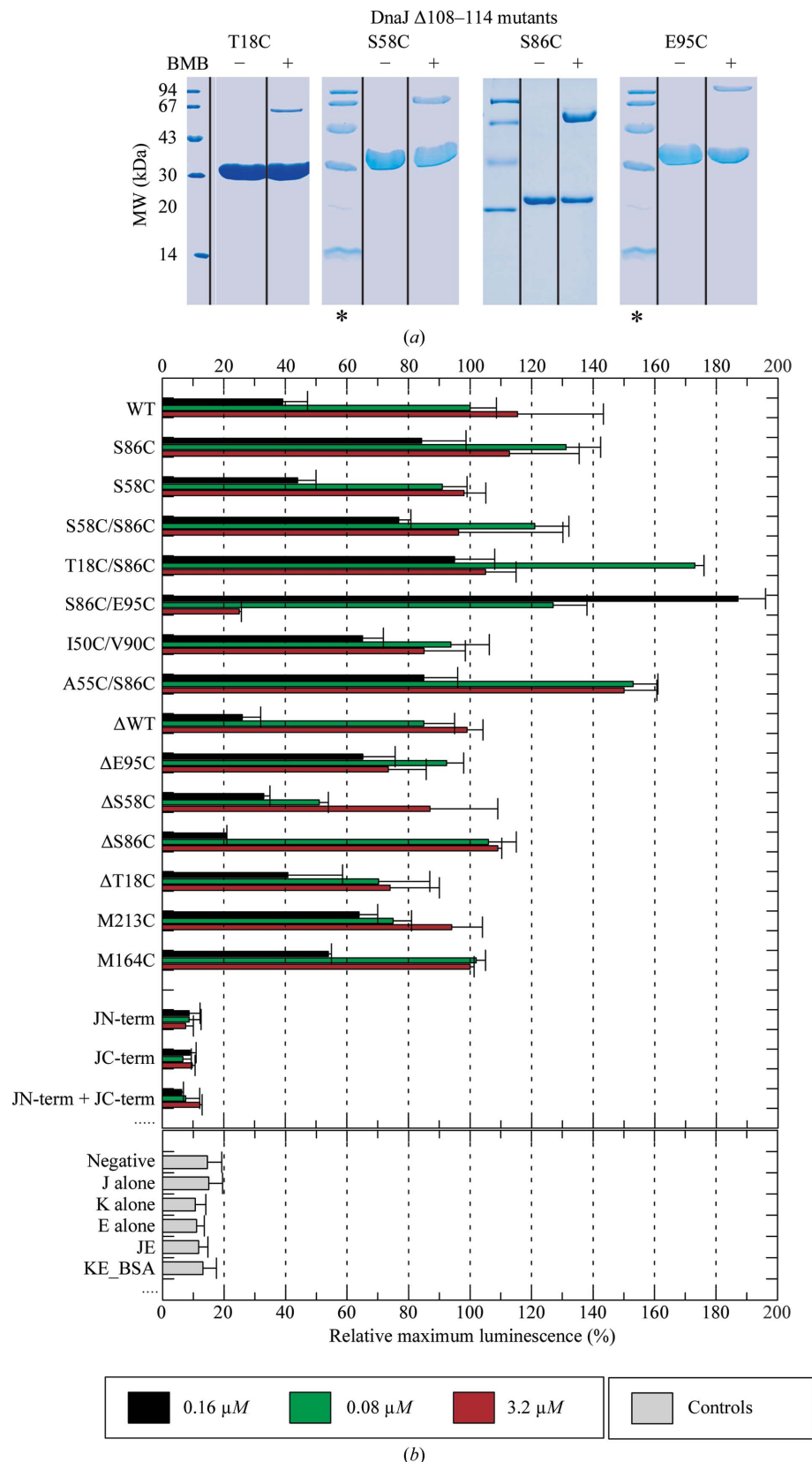
for the J and GF domains to attain a different orientation with respect to the C-terminal domain in each of the four monomers in the asymmetric unit. Analysis with *DynDom* (Hayward & Lee, 2002) shows that these orientations vary by as much as 160° and are related to each other by rotations around the region of amino acids 104–113, which connect the J/GF and C-terminal domains. The lack of electron density for the J and GF domains in the wild-type DnaJ structures indicates an even higher degree of freedom there.

### 3.4. Cross-linking shows that DnaJ is flexible in solution

Given previous findings of predominantly disordered GF domains (Szyperski *et al.*, 1994; Pellicchia *et al.*, 1996; Huang, Ghose *et al.*, 1999) and our observations of a highly ordered GF domain in crystal structures, we initially surmised an equilibrium between a fully ordered and a more disordered state in solution. To test this, we performed cross-linking with the short (~11 Å) bifunctional Cys–Cys cross-linker bismaleimidobutane (Fig. 4*a*) between single cysteine mutations placed in the J/GF domains of DnaJ<sub>Δ108–114</sub>. This resulted in covalently cross-linked DnaJ<sub>Δ108–114</sub> dimers, showing that in solution the J/GF domains within a DnaJ<sub>Δ108–114</sub> dimer can approach each other to within 11 Å, which, given the structure of DnaJ<sub>Δ108–114</sub>, would seem to be possible only when significant disorder exists in solution.

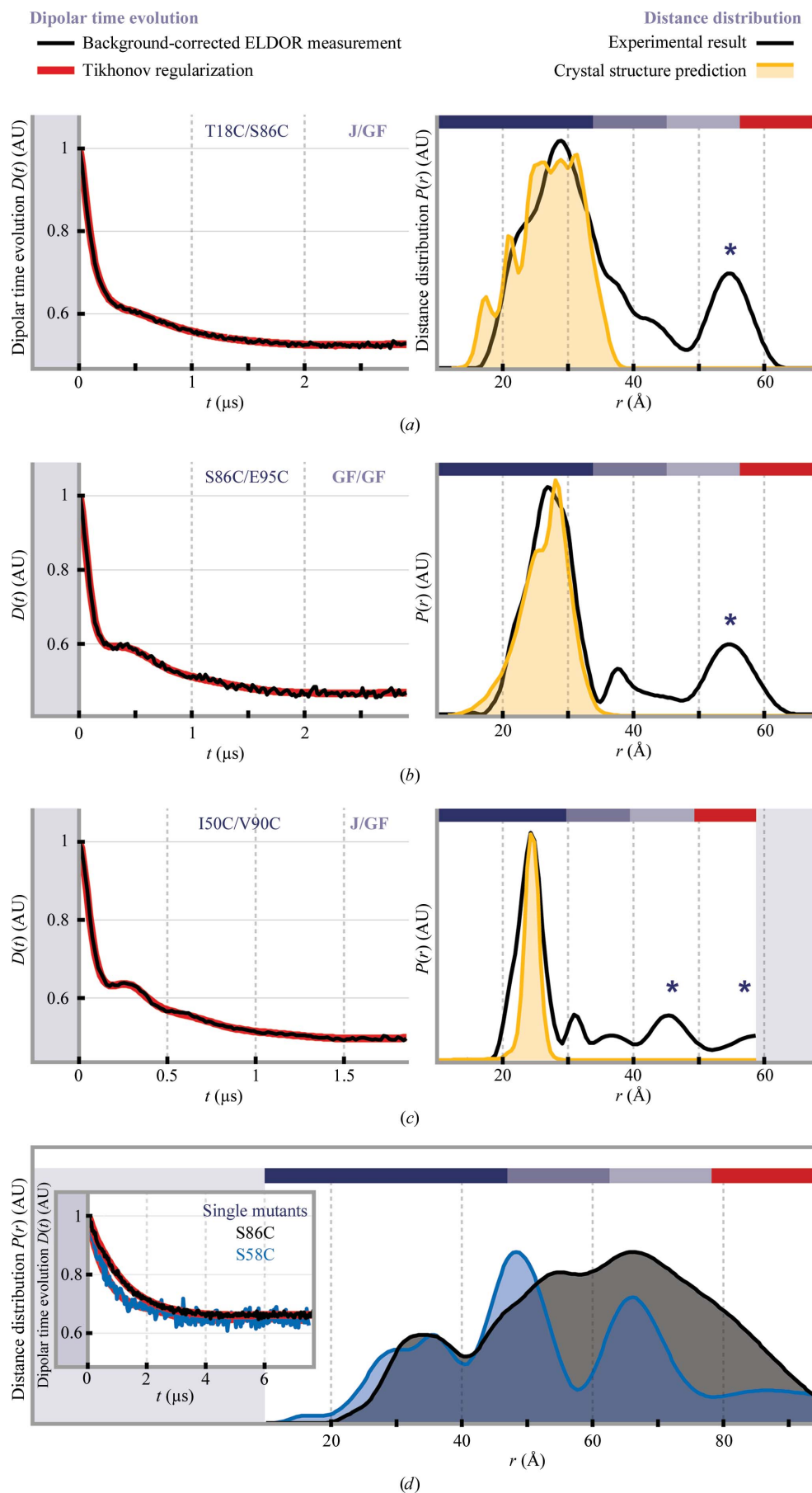
### 3.5. EPR spin-labelling studies show that the GF domain is ordered in solution

To check whether this disorder in solution occurs in the GF domain, four-pulse ELDOR (pELDOR) distance measurements (Martin *et al.*, 1998) were performed on flash-frozen solutions of DnaJ double mutants carrying two spin labels per monomer attached to site-specifically inserted cysteines within the individual J/GF domains (at positions 18 and 86, *i.e.* on the J domain and on the GF domain; at positions 86 and 95, *i.e.* both on the GF domain; and at positions 50 and 90, *i.e.* on the J domain



**Figure 4**  
*(a)* Lanes from SDS–PAGE of cysteine mutants of DnaJ<sub>Δ108–114</sub> without (–) and with (+) cross-linking by BMB after the removal of tetramers by gel filtration (note that the lanes marked \* are the same marker lane shown twice for clarity). *(b)* Luciferase-refolding activity of DnaK assisted by DnaJ. DnaJ variants with mutations in the GF domain show particularly strong modulation of refolding activity. Experiments were conducted in triplicate.





and on the GF domain). Strikingly, these revealed intramonomer distances (in the 20–30  $\text{\AA}$  range) which excellently match the distances derived from the crystal structure (Figs. 5*a–c*). The large modulation depth in the first 250 ns of the time traces governed by these short distances shows that in solution the vast majority of DnaJ molecules adopt the same J/GF structure as observed in the crystal structure. Appendix B shows additional room-temperature X-band continuous-wave (cw) measurements that further support this notion.

However, additional pELDOR signal contributions occur in the 40–60  $\text{\AA}$  range (marked by asterisks in Figs. 5*a–c*) which were not expected from the crystal structure, where intermonomeric distances are only observed at 80  $\text{\AA}$  and above owing to the pronounced V shape of the dimer. Still, these distances indeed stem from intermonomer dipolar interaction as they are also observed in control experiments on two singly labelled DnaJ variants carrying only one spin label per monomer (Fig. 5*d*; labels either at position 58 on the J domain or at position 86 on the GF domain). The broad distance

**Figure 5**  
 (a–c) X-band pELDOR measurements of DnaJ double mutants (a) T18C/S86C, (b) S86C/E95C and (c) I50C/V90C: background-corrected time traces and their Tikhonov regularization fits (left; regularization parameter  $\alpha = 10$ ) and corresponding distance distributions superimposed on their respective rotamer distribution predicted from the crystal structure (right). (d) X-band pELDOR measurements of DnaJ single mutants S86C (black) and S58C (blue) using a long time window for measurements of large distances. Dipolar evolution functions with Tikhonov regularization fit (inset;  $\alpha = 100$ ) and corresponding distance distributions (right). Colour codes above the distributions signify reliability intervals; the reliability of the mean distance of the calculated distribution peaks is shown from dark blue (highest reliability) to red (lowest reliability).

distribution over the whole experimental range shows tails (with low probability down to 20 Å), which is consistent with the cross-linking results and confirms high structural variability in parts of the DnaJ protein backbone in solution.

Given that the individual J/GF sections are structurally rigid at least between Thr18 and Glu95 and that amino acids from Arg118 onwards are clearly visible in the electron-density maps of the wild-type DnaJ crystal structure, any flexibility will most probably be located between Glu95 and Arg118.

### 3.6. Flexibility between the GF and C-terminal domains explains the solution data

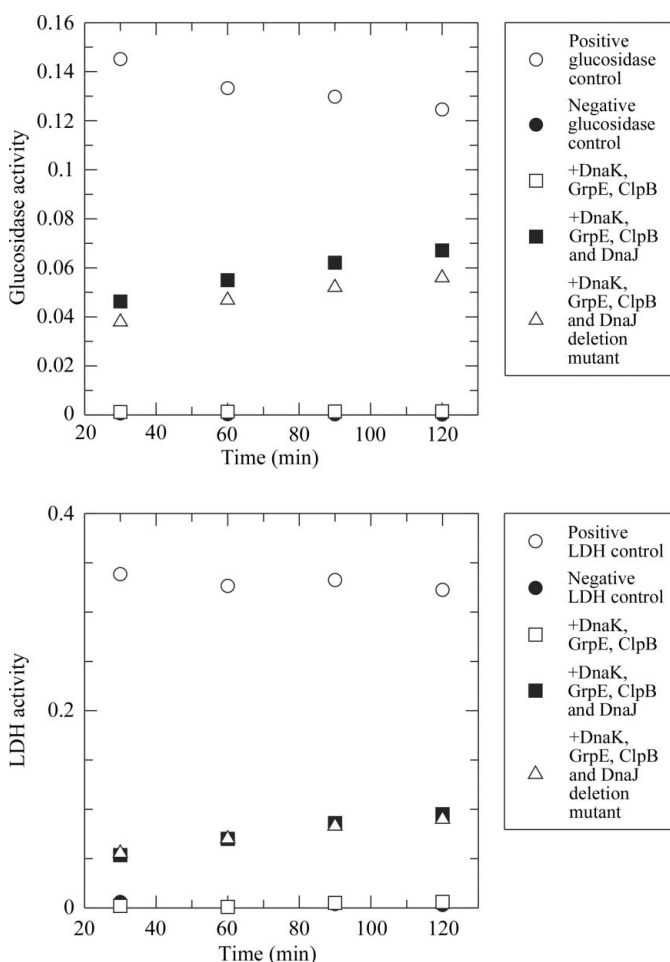
To visualize how much flexibility would be needed to account for the observations described above, we prepared simple geometric models of spin-labelled DnaJ $_{\Delta 108-114}$  *in silico* from the crystal structure. Starting from nearly static models, in which only rotations around the C $^{\alpha}$ –C $^{\beta}$  bonds of the labelled residues were allowed, we progressively allowed rotations around more and more backbone bonds in the connection between the J/GF and C-terminal domains. For each conformation, a pseudo-energy was calculated from (i) the sum of all C $^{\alpha}$ –C $^{\alpha}$  distances smaller than 5 Å and (ii) a term calculated from the EPR distance distribution using the

Boltzmann distribution formula. Using these pseudo-energies, Monte Carlo simulations were performed to obtain ensembles of plausible conformations for all models, which were then checked for their consistency with the cross-linking results by checking for the occurrence of conformations in which the cross-linking sites were within 12 Å of each other during a Monte Carlo run of 20 000 trial structures.

The calculations (see Appendix C) showed that in DnaJ $_{\Delta 108-114}$ , despite the obvious reduction in interdomain flexibility caused by the deletion of seven residues in the J/GF linker, rotation around the backbone bonds of residues 103–110 alone, at the hinge between the J/GF and C-terminal domains, already suffices to explain the cross-linking results.

### 3.7. Mutations in the GF domain modulate the luciferase-refolding activity of DnaK–DnaJ

The various mutations produced for cross-linking experiments were also evaluated for their ability to assist in luciferase refolding by DnaK. Importantly, not only are all of the mutants employed still active, they also display significant modulation of the luciferase refolding activity, in particular when the mutations are located in the GF domain (Fig. 4b).



**Figure 6**  
Refolding assays show that SeMet DnaJ $_{\Delta 108-114/4M}$  is functional.

## 4. Discussion

Given the structural information available previously, the GF domains of DnaJs have frequently been seen as highly disordered flexible structures. The structures and EPR distance distributions of DnaJ presented here necessitate a departure from this view, as they show that in *T. thermophilus* DnaJ the GF domain adopts a highly ordered structure both in the crystal and in solution. Nonetheless, *T. thermophilus* DnaJ is a flexible molecule as shown by cross-linking and EPR studies as well as by the lack of density for the J and GF domains in the DnaJ $_{wt}$  crystal structure. Since the J/GF domains visible in the DnaJ $_{\Delta 108-114}$  crystal structure each assume a different orientation relative to their respective C-terminal domains, it seems likely that this flexibility is, at least to a large extent, concentrated at the junction between the J/GF domain and the C-terminal domain. Indeed, our very simple molecular modelling studies showed that flexibility in this region alone is sufficient to explain the EPR and cross-linking data. Possibly, the effects that mutations in the GF region have on luciferase refolding are caused by a modulation of this flexibility. In support of this hypothesis, one may compare the *T. thermophilus* and *E. coli* DnaK–DnaJ systems. In *E. coli* DnaJ, which stimulates DnaK ATPase activity strongly, a highly flexible GF domain was observed in NMR studies, although a partial structure was observed (Pellecchia *et al.*, 1996), and interactions with helix III of the J domain (as observed in *T. thermophilus* DnaJ in the current study) were surmised (Huang, Flanagan *et al.*, 1999). In contrast, in *T. thermophilus* DnaJ, which stimulates DnaK ATPase activity only weakly, the structures presented here indicate a strongly ordered GF domain.

A possible explanation could lie in the structure of the DnaK–DnaJ complex. If, as suggested (Han & Christen, 2003; Zuiderweg & Ahmad, 2012), the DnaJ C-terminal domain binds to substrate polypeptide which is also bound to the peptide-binding domain of DnaK, the DnaJ J domain must reach over to another position (Jiang *et al.*, 2007; Ahmad *et al.*, 2011; Sousa *et al.*, 2012; Zuiderweg & Ahmad, 2012) to effect ATPase stimulation, a situation in which flexibility between the two DnaJ domains would be beneficial and might be modulated for regulatory purposes.

### APPENDIX A Kinetic measurements

The functionality of SeMet DnaJ<sub>Δ108–114/4M</sub> was ascertained by refolding assays with lactate dehydrogenase and  $\alpha$ -glucosidase as follows: 0.2  $\mu$ M LDH was denatured for 30 min in a water bath heated to 353 K in denaturation buffer (50 mM MOPS/NaOH pH 7.5, 150 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM ATP, 2 mM DTE). For refolding, 192  $\mu$ l of the denatured sample was added to 8  $\mu$ l chaperone mixture (final concentrations of 1  $\mu$ M ClpB, 1.6  $\mu$ M DnaK, 0.4  $\mu$ M DnaJ or DnaJ<sub>Δ108–114</sub>, 0.2  $\mu$ M GrpE in denaturation buffer) and incubated for 30, 60, 90 and 120 min. After these refolding periods, 20  $\mu$ l samples were mixed with 180  $\mu$ l 50 mM Tris–HCl pH 7.5, 50 mM KCl, 250  $\mu$ M NADH, 10 mM pyruvate. The turnover of pyruvate and NADH was measured at 298 K by following the decrease

of the NADH absorption signal at 340 nm using a VarioScan MTP reader. The rate was calculated from the linear range of the curves and was plotted against the refolding time (Fig. 6a).

0.2  $\mu$ M  $\alpha$ -glucosidase in denaturation buffer was denatured for 8 min in a water bath heated to 348 K. For refolding, 192  $\mu$ l of the denatured sample was added to 8  $\mu$ l chaperone mixture (final concentrations of 1  $\mu$ M ClpB, 1.6  $\mu$ M DnaK, 0.4  $\mu$ M DnaJ or DnaJ<sub>Δ108–114</sub>, 0.2  $\mu$ M GrpE in denaturation buffer) and incubated at 328 K for 30, 60, 90 and 120 min. After these refolding periods, 20  $\mu$ l samples were mixed with 180  $\mu$ l 50 mM potassium phosphate pH 6.8, 2 mM PNP- $\alpha$ -GP (*para*-nitrophenyl- $\alpha$ -D-glucopyranoside), 0.1 mg ml<sup>-1</sup> BSA in a Corning microtitre plate to yield a final concentration of 20 nM  $\alpha$ -glucosidase. The turnover of PNP- $\alpha$ -GP was recorded in a VarioScan MTP reader, following the decrease of absorption at 405 nm at 313 K. The rate was calculated from the linear range of the curve and plotted against the incubation time of refolding (Fig. 6b).

### APPENDIX B Room-temperature continuous-wave EPR

All samples examined by pELDOR spectroscopy at cryogenic temperatures (80 and 50 K) were also examined at ambient temperature conditions in an X-band continuous-wave EPR experiment. The width of the observed spectrum can be directly related to a rotational correlation time of the spin-label side chain (*via* the *chili* routine of the *MatLab* package *EasySpin*; Stoll & Schweiger, 2006) and is thus a probe of the immediate surroundings at the label site. Fig. 7 shows the recorded spectra as well as estimates for the E95C and T18C single spectra (marked E95C and T18C) in which a 50% proportion of the S86C spectrum was subtracted from the double mutants S86C/E95C and T18C/S86C. For the simulations, the **g** and **A** matrices necessary for the fit were determined from W-band measurements to be **g** = (2.0086, 2.0064, 2.0023) and **A** = (17, 18, 104) MHz (not shown). The rotational correlation times were determined to be 2–5 ns for the single-label spectra and 5 and 15 ns for the I50C/V90C double-label construct (I50C/V90C exhibits especially slow tumbling and strong anisotropy, but both components are still within the same order of magnitude).

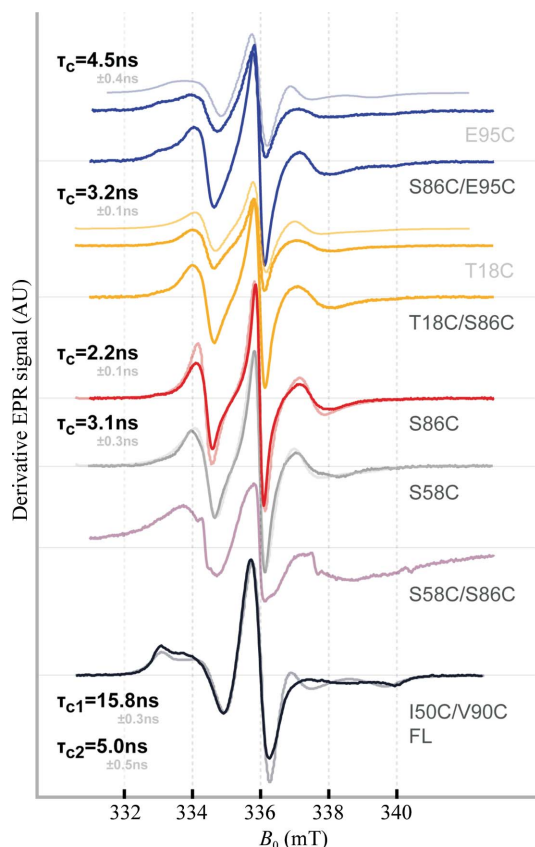


Figure 7 Room-temperature continuous-wave EPR results.

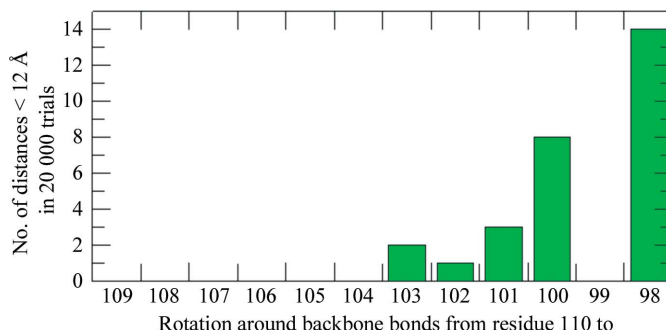


Figure 8 Cross-linking feasibility simulation results.

In an extensive study, Bordignon *et al.* (2005) investigated the whole range of labels completely immobilized in a protein to those attached to side chains in unfolded sections. Even the fastest side-chain motion observed for DnaJ S86C is much slower than that observed for a flexible protein backbone. This is further evidence for the GF domain not being unfolded but in a rigid conformation, as observed by pELDOR and crystallography.

## APPENDIX C

### Computational modelling

To calculate ensembles of DnaJ conformers, EPR distance distributions  $p(r)$  were converted to pseudo-potential energy surfaces using  $E_{\text{pseudo}}^{\text{pot}}(r) = -\ln p(r)$ . To avoid clashes, a penalty term  $w n_{\text{clashes}}$  was added to the pseudo-energies, where  $w$  is a weighting factor and  $n_{\text{clashes}}$  is the number of C $^{\alpha}$  atoms within 3 Å of each other. These pseudo-potential energy surfaces were then used to guide Monte Carlo simulations of DnaJ using specially written *MatLab* scripts, in which only rotations around selected bonds (linker main-chain bonds, see text, and C $^{\alpha}$ –C $^{\beta}$  bonds of labelled residues, which were simulated by tyrosines) were allowed. To check for the possibility of cross-link formation within cysteine mutant dimers, the number of cysteine–cysteine distances smaller than 12 Å was counted in an ensemble of 20 000 structures. Starting with the backbone of residue 110 in DnaJ $_{\Delta 108-114}$ , progressively more backbone bonds were allowed to rotate. When allowing rotations around the backbone bonds of residues 103–110, the first instances of cysteine–cysteine distances below 12 Å were observed. Including more bonds caused a steep increase in the number of distances below 12 Å, as expected (Fig. 8).

We thank Marion Hülseweh and Arne Rufer for technical assistance, Melanie Müller for MALDI–TOF and Ingrid Vetter for crystallographic software support. Diffraction data were collected on beamline X10SA at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland. We thank the Dortmund–Heidelberg data-collection team and the PXII staff for their support in setting up the beamline. We also thank Max Cryle for fruitful discussions and for suggesting the use of a cross-linking agent.

## References

- Adzhubei, A. A. & Sternberg, M. J. (1993). *J. Mol. Biol.* **229**, 472–493.
- Ahmad, A., Bhattacharya, A., McDonald, R. A., Cordes, M., Ellington, B., Bertelsen, E. B. & Zuiderweg, E. R. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 18966–18971.
- Aponte, R. A., Zimmermann, S. & Reinstein, J. (2010). *J. Mol. Biol.* **399**, 154–167.
- Aron, R., Lopez, N., Walter, W., Craig, E. & Johnson, J. (2005). *Genetics*, **169**, 1873–1882.
- Banumathi, S., Zwart, P. H., Ramagopal, U. A., Dauter, M. & Dauter, Z. (2004). *Acta Cryst.* **D60**, 1085–1093.
- Beinker, P., Schlee, S., Groemping, Y., Seidel, R. & Reinstein, J. (2002). *J. Biol. Chem.* **277**, 47160–47166.
- Bordignon, E., Klare, J. P., Doebber, M., Wegener, A. A., Martell, S., Engelhard, M. & Steinhoff, H. J. (2005). *J. Biol. Chem.* **280**, 38767–38775.
- Bork, P., Sander, C., Valencia, A. & Bukau, B. (1992). *Trends Biochem. Sci.* **17**, 129.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Caplan, A. J., Cyr, D. M. & Douglas, M. G. (1993). *Mol. Biol. Cell*, **4**, 555–563.
- Cyr, D. M., Langer, T. & Douglas, M. G. (1994). *Trends Biochem. Sci.* **19**, 176–181.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Doublé, S. (1997). *Methods Enzymol.* **276**, 523–530.
- Dumitru, G. L., Groemping, Y., Klostermeier, D., Restle, T., Deuerling, E. & Reinstein, J. (2004). *J. Mol. Biol.* **339**, 1179–1189.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Groemping, Y., Klostermeier, D., Herrmann, C., Veit, T., Seidel, R. & Reinstein, J. (2001). *J. Mol. Biol.* **305**, 1173–1183.
- Groemping, Y., Seidel, R. & Reinstein, J. (2005). *FEBS Lett.* **579**, 5713–5717.
- Han, W. & Christen, P. (2003). *J. Biol. Chem.* **278**, 19038–19043.
- Hayward, S. & Lee, R. A. (2002). *J. Mol. Graph. Model.* **21**, 181–183.
- Hendrick, J. P. & Hartl, F. U. (1993). *Ann. Rev. Biochem.* **62**, 349–384.
- Hendrick, J. P., Langer, T., Davis, T. A., Hartl, F. U. & Wiedmann, M. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 10216–10220.
- Hu, J., Wu, Y., Li, J., Qian, X., Fu, Z. & Sha, B. (2008). *BMC Struct. Biol.* **8**, 3.
- Huang, K., Flanagan, J. M. & Prestegard, J. H. (1999). *Protein Sci.* **8**, 203–214.
- Huang, K., Ghose, R., Flanagan, J. M. & Prestegard, J. H. (1999). *Biochemistry*, **38**, 10567–10577.
- Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D. & Jung, H. (2006). *Appl. Magn. Reson.* **30**, 473–498.
- Jeschke, G. & Polyhach, Y. (2007). *Phys. Chem. Chem. Phys.* **9**, 1895–1910.
- Jiang, J., Maes, E. G., Taylor, A. B., Wang, L., Hinck, A. P., Lafer, E. M. & Sousa, R. (2007). *Mol. Cell*, **28**, 422–433.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Karzai, A. W. & McMacken, R. (1996). *J. Biol. Chem.* **271**, 11236–11246.
- Klostermeier, D., Seidel, R. & Reinstein, J. (1999). *J. Mol. Biol.* **287**, 511–525.
- Laufen, T., Mayer, M. P., Beisel, C., Klostermeier, D., Mogk, A., Reinstein, J. & Bukau, B. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 5452–5457.
- Li, J., Wu, Y., Qian, X. & Sha, B. (2006). *Biochem. J.* **398**, 353–360.
- Martin, R. E., Pannier, M., Diederich, F., Gramlich, V., Hubrich, M. & Spiess, H. W. (1998). *Angew. Chem. Int. Ed.* **37**, 2834–2837.
- McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* **D61**, 458–464.
- McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
- Motohashi, K., Taguchi, H., Ishii, N. & Yoshida, M. (1994). *J. Biol. Chem.* **269**, 27074–27079.
- Motohashi, K., Yohda, M., Endo, I. & Yoshida, M. (1996). *J. Biol. Chem.* **271**, 17343–17348.
- Motohashi, K., Yohda, M., Odaka, M. & Yoshida, M. (1997). *FEBS Lett.* **412**, 633–636.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. SERC Daresbury Laboratory, Warrington, UK.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pannier, M., Veit, S., Godt, A., Jeschke, G. & Spiess, H. W. (2000). *J. Magn. Reson.* **142**, 331–340.

- Pellecchia, M., Szyperski, T., Wall, D., Georgopoulos, C. & Wüthrich, K. (1996). *J. Mol. Biol.* **260**, 236–250.
- Pierpaoli, E. V., Sandmeier, E., Schönfeld, H. J. & Christen, P. (1998). *J. Biol. Chem.* **273**, 6643–6649.
- Polyhach, Y., Bordignon, E. & Jeschke, G. (2011). *Phys. Chem. Chem. Phys.* **13**, 2356–2366.
- Popp, S. L. & Reinstein, J. (2009). *FEBS Lett.* **583**, 573–578.
- Qiu, X.-B., Shao, Y.-M., Miao, S. & Wang, L. (2006). *Cell. Mol. Life Sci.* **63**, 2560–2570.
- Ramagopal, U. A., Dauter, Z., Thirumuruhan, R., Fedorov, E. & Almo, S. C. (2005). *Acta Cryst.* **D61**, 1289–1298.
- Rassow, J., Voos, W. & Pfanner, N. (1995). *Trends Cell Biol.* **5**, 207–212.
- Ravelli, R. B. G., Leiros, H.-K. S., Pan, B., Caffrey, M. & McSweeney, S. (2003). *Structure*, **11**, 217–224.
- Read, R. J. (1986). *Acta Cryst.* **A42**, 140–149.
- Russell, R., Karzai, A. W., Mehl, A. F. & McMacken, R. (1999). *Biochemistry*, **38**, 4165–4176.
- Schlee, S. & Reinstein, J. (2002). *Cell. Mol. Life Sci.* **59**, 1598–1606.
- Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* **D58**, 1772–1779.
- Sha, B., Lee, S. & Cyr, D. M. (2000). *Struct. Fold. Des.* **8**, 799–807.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Sousa, R., Jiang, J., Lafer, E. M., Hinck, A. P., Wang, L., Taylor, A. B. & Maes, E. G. (2012). *Proc. Natl Acad. Sci. USA*, **109**, E734.
- Stoll, S. & Schweiger, A. (2006). *J. Magn. Res.* **178**, 42–55.
- Suzuki, H., Noguchi, S., Arakawa, H., Tokida, T., Hashimoto, M. & Satow, Y. (2010). *Biochemistry*, **49**, 8577–8584.
- Svärd, M., Biterova, E., Bourhis, J.-M. & Guy, J. (2011). *PLoS One*, **6**, e22337.
- Szyperski, T., Pellecchia, M., Wall, D., Georgopoulos, C. & Wüthrich, K. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11343–11347.
- Terwilliger, T. (2003). *Methods Enzymol.* **374**, 22–37.
- Terwilliger, T. (2004). *J. Synchrotron Rad.* **11**, 49–52.
- Tikhonov, A. N. & Arsenin, Y. (1977). *Solutions of Ill-posed Problems*. New York: Wiley.
- Todd, A. P., Cong, J., Levinthal, F., Levinthal, C. & Hubell, W. L. (1989). *Proteins*, **6**, 294–305.
- Todd, A. P., Crozel, V., Levinthal, F., Levinthal, C. & Hubbell, W. (1987). *Biophys. J.* **51**, A83.
- Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.
- Wall, D., Zyllicz, M. & Georgopoulos, C. (1994). *J. Biol. Chem.* **269**, 5446–5451.
- Wall, D., Zyllicz, M. & Georgopoulos, C. (1995). *J. Biol. Chem.* **270**, 2139–2144.
- Walsh, P., Bursać, D., Law, Y. C., Cyr, D. & Lithgow, T. (2004). *EMBO Rep.* **5**, 567–571.
- Watanabe, Y. H. & Yoshida, M. (2004). *J. Biol. Chem.* **279**, 15723–15727.
- Yan, W. & Craig, E. A. (1999). *Mol. Cell. Biol.* **19**, 7751–7758.
- Zuiderweg, E. R. P. & Ahmad, A. (2012). *Proc. Natl Acad. Sci. USA*, **109**, E735.
- Zwart, P. H., Banumathi, S., Dauter, M. & Dauter, Z. (2004). *Acta Cryst.* **D60**, 1958–1963.