Structural basis of the allosteric trigger of the Hsp70 chaperone proteins.

Akash Bhattacharya, Alexander V. Kurochkin, Grover N.B. Yip, Eric B. Bertelsen and Erik R. P. Zuiderweg

Department of Biological Chemistry and Biophysics Research Division, The University of Michigan.

This work solves a decades-old dilemma that stood in the way of understanding the allosteric mechanism of Hsp70 (heat shock 70 kDa) chaperone proteins. Hsp70s are central to protein folding, refolding, and trafficking in organisms ranging from *Archae* to *Homo Sapiens*, both at normal and at stressed cellular conditions[1]. Hsp70s are comprised of two main domains: a 44 kDa N-terminal nucleotide-binding domain (NBD), and a 25 kDa substrate-binding domain (SBD) that harbors the substrate binding site. The nucleotide binding site in the NBD and the substrate binding site in the SBD are allosterically linked: ADP binding promotes substrate binding, while ATP binding promotes substrate release[2] [3]. It has long been a goal of structural biology to characterize the nature of the allosteric coupling in these proteins. However, even the most sophisticated X-ray crystallography studies of the isolated NBD could show no difference in overall conformation between the ATP and ADP state [4,5]. Hence the dilemma: how is the state of the nucleotide communicated between NBD and SBD? The solution of the dilemma is especially interesting in light of the fact that Hsp70s are ancient proteins, and amongst the first allosteric proteins in nature [6].

Here we report a solution NMR study of the NBD of the Hsp70 from *Thermus thermophilus*, in the APO, ADP and AMP-PNP states, where the latter is a non-hydrolysable ATP analogue. Using the modern NMR methods of residual dipolar coupling analysis[7] [8], we discovered that the nucleotide binding cleft opens up by as much as 20 degrees between the AMP-PNP (closed) and ADP (open) state. We also discover that a surface cleft, hypothesized to be essential for the allosteric coupling between NBD and SBD[9] , echoes these changes. Hence, the nature of the allosteric trigger and coupling for Hsp70 chaperones is revealed here for the first time, solving the dilemma.

Recently, Hsp70s have been linked to cancer and diseases associated with protein misfolding such as Alzheimer's[10], Parkinson's[11] , and Huntington's[12] [13]. It has been suggested

that modulation of Hsp70 activity with small compounds may form an avenue to treat these diseases[14]. It is particularly attractive to try to interfere with the Hsp70 allosteric mechanism, since this would provide intrinsic selectivity. One such target may constitute the NBD allosteric trigger.

In this work, we study the NBD of the Hsp70 of *Thermos Thermophilus*[15]. Because of the very high homology between this protein and the 11 Hsp70 gene products in humans, we expect the derived conclusions of our work to be valid for the human proteins as well. The NBD's of the Hsp70s are comprised of four subdomains[4] called IA, IB, IIA and IIB, as defined in Figure 1. The nucleotide, ATP or ADP.PO4, sits deeply in the central cleft interacting non-covalently with all four subdomains. An overlay of five available crystal structures of the NBD of bovine Hsc70 in APO, ADP (twice), ADP-V_2O_5 state and the mutant K71M complexed with ATP, shows no significant differences between these states (see Figure 1). The results of the current solution NMR study of the conformational changes of the NBD of the Hsp70 from *Thermos Thermophilus*, hereafter referred to as DnaK-Tth, in the APO, ADP and AMP-PNP state, are summarized in Figure 2. In contrast to the results shown in Figure 1, it is evident from Figure 2 that major differences exist between the different ligand states in solution. The largest changes occur for domain IIB (right top). In the AMP-PNP state, the nucleotide binding cleft is mostly closed; in the APO state it is mostly opened, and the ADP state resembles the APO state.

The conformational changes were determined using residual dipolar coupling NMR (RDC) analysis[8]. Briefly, the protein was dissolved in a dilute liquid crystal[16] providing an oriented environment for the protein in which the magnetic dipolar coupling of the amide nitrogen and hydrogen nuclei can be measured by NMR techniques[17 18]. The magnitudes and signs of these couplings report on the orientations of the protein's NH inter-nuclear vectors with respect to the magnetic field [17]. This information is, in turn, used to obtain the orientation of the protein's sub-domains with respect to the magnetic field, making use of the available high-resolution (X-ray) structures of the sub-domains. Subsequently, this information is used to reconstruct the relative orientation of the sub-domains. Sub-domain orientation is characterized by three axes corresponding to a rectangular parallelepiped (matchbox-shape) with the longest axis defined as S_{zz}, the shortest as S_{xx} and S_{yy} in between.

Figure 3 shows the quality of the orientational data for the three nucleotide states. In these globe plots, the main orientational axis S_{zz} for all subdomains is located approximately at 40 ^{o}West, 5 ^{o}South (in the red circle). This average location by itself is irrelevant, as it just reports the difference between the physical S_{zz} axis orientation of the aligned protein and the

arbitrary PDB-file Z-axis direction of the model structure used. However, the differences between the globe locations of the S_{zz} axes of the different subdomains are highly relevant: they indicate that some of the subdomains are oriented differently than in the model structure. The spread in the S_{zz} directions of the subdomains represents the uncertainty in the raw data.

Accordingly, we conclude that all subdomains in the AMP-PNP state are, within experimental error, similarly oriented as in the model structure, which is based on an Hsc70 NBD crystal structure in the ADP state (3HSC.pdb, ref [4]). In contrast, in the APO and ADP states, one observes statistically significant deviations, by as much as 20 ^{o}, for the S_{zz} axes of subdomains IIA and IIB from those of IA and IB. Figure 3 also shows that the orientations of the S_{xx} (and S_{yy}) axes of the sub-domains are insufficiently defined for all states to draw reliable conclusions about these directions (see legend). The S_{zz}-orientational data was translated to the coordinate models in Figure 2 by superposing each subdomain on the coordinates of the sub-domains of the model structure, which in turn was rotated to the average solution orientation. In the superposition, the Ca atom coordinates of the secondary structure elements in the moveable subdomains were optimized with respect to those of the reference structure by translation and by re-orientation around the S_{zz} axes of the subdomains only. That is, the obtained S_{zz} orientations were kept fixed, while allowing the S_{xx} (and S_{yy}) orientations to be optimized. Finally, the superposition of all three structures in Figure 2 was obtained by optimizing the overlay of subdomains IA and IB of each structure.

Figure 2 shows that the cleft opens by as much as 20 degrees between AMP-PNP and ADP.Pi states. We interpret this finding as follows. In the AMP-PNP state, the left and right halves of the protein are bridged by the rigid nucleotide mimic. Such should also be the case if ATP is the ligand. For ADP.Pi, however, the γ-phosphate bond is hydrolyzed, which breaks the molecule and thus the bridge between left and right, allowing an opening of the cleft. Likely, the left and right halves can move relatively independently in the APO and ADP.Pi states. Potentially, the ADP.Pi and APO structures can open dynamically even further than displayed in the models. The current findings complement previous NMR study[19] where it was shown that the ATP state, and especially the ADP.AlF_x state, of DnaK-Tth NBD are characterized by a single conformation, while at least two conformations, in a slow, but dynamic equilibrium, exist for the ADP.Pi state. It is also relevant to cite the chemical shift differences reported for the different nucleotide states of the protein in Figure 4 of that work[19]. It shows that most of the chemical shift changes between ADP.Pi and ADP.AlF_x states refer to the inside of the

nucleotide-binding cleft, also far away from the nucleotide binding site. This is of course well-explained by the here-detected global change in subdomain orientation accompanying such nucleotide changes.

In addition, major chemical shift changes on the interface between domains IA and IIA upon the nucleotide exchange have been reported[19]. These changes are now explained by the statistically significant rotation of subdomain IIA between the different states, as disclosed in Figure 3. IIA is seen to rotate in opposite direction of subdomain IIB. Hence, the combined motion can be seen as a scissor motion. In our modeling, the hinge of the scissor runs in the middle of the beta sheet (see white residues in Figure 2). However it should be kept in mind that the precise hinge position is somewhat dependent on the details of the superpositioning method.

Figure 4 shows that the surface cleft running between domains IA and IIA ("bottom") opens up in the ADP and APO states. The color coding in Figure 4 shows residues that, when mutated[20 21] , affect the allosteric coupling between NBD and SBD. It also shows residues for which chemical shift changes could be identified between ADP and ATP state in *E.coli* DnaK, which are attributed to a changing interaction with the conserved linker between NBD and SBD[9]. It is clearly seen that these residues are all in the surface cleft area, and that the accessibility of several of them is affected by the rotations of domain IIA. This strongly supports the recent suggestions that the cleft is an allosteric binding site for the linker[9 22]. In the ADP state, the linker moves freely and allows the SBD to move relatively freely in solution (Ref [9], and Bertelsen and Zuiderweg, submitted). In the ATP state, the linker is immobilized, likely in this cleft[9]. From our data and the modeling in Figure 4, we conclude, indeed, that the cleft changes its shape between the states, which could account for the changing binding mode of the linker.

As shown in Fig 1, all available X-ray structures of the isolated NBD correspond closely to each other regardless their nucleotide state. As it turns out, they all correspond to the closed state in solution as shown in Fig 5a. It is likely that crystal packing promotes the closed state no matter the nature of the nucleotide. This indicates that there is very little or no energy, that resists the closing of the open state of the (isolated) NBD. Figure 5b shows a superposition of the open state (ADP.Pi) with the *E.coli* DnaK NBD complexed with the nucleotide exchange factor GrpE[23]. The correspondence indicates, in contrast to earlier proposals where GrpE was thought to actively "break open" the nucleotide binding cleft, that this open state is already available in the ADP form, even without bound GrpE. Hence, the older mechanism, suggesting that GrpE forces an induced fit which actively promotes nucleotide exchange[24], should be

modified to incorporate that GrpE selectively captures and stabilizes the ADP state as compared to the closed, ATP state. The latter mechanism, selective capture, has recently been recognized for several allosteric proteins[25] and nucleic acids[26]. The nucleotide exchange process in the Hsp70 chaperones (abundant ATP replacing ADP) in this model is then catalyzed by other processes, which may include the interaction of GrpE's tail with the SBD, as has been suggested before[23].

Methodological Summary

DnaK-Tth 1–381 (DnaK-Tth NBD) with N-terminal His-tag was expressed in *E.coli* BL21 cells grown in a triple-labeled M-9 medium containing ^{15}N-ammonium chloride. The protein was purified in a standard way with affinity chromatography on the Ni-NTA agarose and was eluted from a FastFlo Q-column with Imidazole gradient [27].

NMR samples contained 0.3-0.4 mM protein in 50mM HEPES, pH7.4, 10mM KCl, 5mM $MgCl_2$, and 5mM sodium phosphate. ADP or AMP-PNP concentration was 10 mM. Experiments were performed at 50 ^{o}C on an 800 MHz Varian Inova spectrometer, using a triple resonance cold-probe. Backbone resonance assignments were obtained from a single 3D HNCA-TROSY experiment for each nucleotide state, using a previously obtained peak list for DnaK-Tth in the ADP-AlF_x state as a template[28]. 223, 281 and 310 assignments were obtained for the APO, ADP and AMP-PNP form, respectively.

For RDC measurements, Pf1 bacteriophage in the aforementioned buffer was added to the NMR samples to the final concentration of 20 mg/ml for partial alignment[29]. RDCs were extracted from a series of 2D TROSY experiments with a [κt1/2-180(N,H)-κt1/2] sequence[30] at the beginning of the ^{15}N chemical shift labeling period , with κ=0, 0.75 and 1.5. Table 1 gives the number of RDCs obtained per subdomain, for each state of the protein.

The RDC data was used to orient subdomains of a DnaK-Tth homology model, which was based on aligning sequence and secondary structure elements of DnaK-Tth to the crystal structure of the NBD of bovine Hsc70 in the ADP state[4] (3HSC.pdb). It was assumed that each subdomain was a rigid unit by itself. We used REDCAT[31] (A Residual Dipolar Coupling Analysis Tool) to transform the RDCs to the orientational data. REDCAT's solution algorithm relies on

singular value decomposition and Monte Carlo sampling to generate a best-fit solution consistent with the input RDCs. The computations yielded an ensemble of 1000 structures compatible with the input structure and the set of RDCs provided, based on an experimental error range of 3 Hz. The validity of the error estimation was confirmed using a self-validation procedure.

References

1 B. Bukau and A. L. Horwich, *Cell* **92** (3), 351 (1998).
2 M. P. Mayer, S. Rudiger, and B. Bukau, *Biol Chem* **381** (9-10), 877 (2000).
3 J. S. McCarty, A. Buchberger, J. Reinstein et al., *J. Mol. Biol.* **249** (1), 126 (1995).
4 K. M. Flaherty, C. Deluca-Flaherty, and D. B. McKay, *Nature* **346** (6285), 623 (1990).
5 M. C. O'Brien, K. M. Flaherty, and D. B. McKay, *J. Biol. Chem.* **271** (27), 15874 (1996).
6 R.S. Gupta, *Microbiol Mol Biol Review* **62**, 1435 (1998).
7 J. R. Tolman, J. M. Flanagan, M. A. Kennedy et al., *Proc. Natl. Acad. Sci. USA* **92** (20), 9279 (1995).
8 M.W. Fischer, J.A. Losonczi, J.L. Weaver et al., *Biochemistry* **38** (28), 9013 (1999).
9 J. F. Swain, G. Dinler, R. Sivendran et al., *Mol Cell* **26** (1), 27 (2007).
10 J. E. Gestwicki, G. R. Crabtree, and I. A. Graef, *Science* **306** (5697), 865 (2004).
11 K.K. Chung and T.M. Dawson, *Neuron* **44**, 899 (2004).
12 S. Krobitsch and S. Lindquist, *Proc Natl Acad Sci U S A* **97** (4), 1589 (2000).
13 T.V. Novoselova, B.A. Margulis, S.S. Novoselov et al., *J. Neurochemistry* **94**, 597 (2005).
14 J. L. Brodsky and G. Chiosis, *Curr Top Med Chem* **6** (11), 1215 (2006).
15 D. Klostermeier, R. Seidel, and J. Reinstein, *J Mol Biol* **279** (4), 841 (1998).
16 N. Tjandra, J. G. Omichinski, A. M. Gronenborn et al., *Nat. Struct. Biol.* **4** (9), 732 (1997).
17 J. R. Tolman, H. M. Al-Hashimi, L. E. Kay et al., *J. Am. Chem. Soc.* **123** (7), 1416 (2001).
18 D. Yang, J.R. Tolman, N.T. Goto et al., *J. Biomol. NMR* **12** (2), 325 (1998).
19 M. Revington, T. M. Holder, and E. R. Zuiderweg, *J Biol Chem* **279** (32), 33958 (2004).
20 M. Vogel, B. Bukau, and M. P. Mayer, *Mol Cell* **21** (3), 359 (2006).
21 M. P. Mayer, T. Laufen, K. Paal et al., *J Mol Biol* **289** (4), 1131 (1999).
22 M. Vogel, M. P. Mayer, and B. Bukau, *J Biol Chem* **281** (50), 38705 (2006).
23 C.J. Harrison, M. Hayer-Hartl, M. Di Liberto et al., *Science* **276**, 431 (1997).
24 C. J. Harrison, M. Hayer-Hartl, M. Di Liberto et al., *Science* **276** (5311), 431 (1997).
25 D. Kern and E. R. Zuiderweg, *Curr Opin Struct Biol* **13** (6), 748 (2003).
26 H. M. Al-Hashimi, Y. Gosser, A. Gorin et al., *J Mol Biol* **315** (2), 95 (2002).
27 M. Revington, Y. Zhang, G. N. Yip et al., *J Mol Biol* **349** (1), 163 (2005).
28 M. Revington and E. R. Zuiderweg, *J Biomol NMR* **30** (1), 113 (2004).
29 M.R. Hansen, L. Mueller, and A. Pardi, *Nat. Struct. Biol.* **5** (12), 1065 (1998).
30 D. Yang, R.A. Venters, G.A. Mueller et al., *J. Biomol. NMR* **14** (4), 333 (1999).
31 H. Valafar and J. H. Prestegard, *J Magn Reson* **167** (2), 228 (2004).
32 C. S. Gassler, A. Buchberger, T. Laufen et al., *Proc Natl Acad Sci U S A* **95** (26), 15229 (1998).

Acknowledgements: We thank the W. F. Keck Foundation, National Science Foundation (NSF) and National Institutes of Health (NIH) for funds for the purchase of an 800MHz spectrometer and cryogenic probe. The research was supported by an NIH grant.

Author Contributions:
A.B. and A.V.K collected the NMR data. A.B. analyzed the data, with initial help from G.N.B.Y. and E.B.B. and E.R.P.Z. A.B. wrote parts of the manuscript. A.V.K prepared the protein samples. E.R.P.Z. conceived the project and NMR methods and wrote parts of the manuscript.

Author Information:
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to E.R.P.Z. (zuiderwe@umich.edu).

Table 1: RDCs per subdomain			
Subdomain	Protein state		
	APO	ADP	AMPPNP
IA	66	66	51
IB	32	38	36
IIA	25	43	41
IIB	30	50	38

Table 2. Superposition RMSD				
	Superposition	TTh APO	Tth ADP	Tth AMP-PNP
HSC-bovine ADP	IA IB Best[a] (Å)	1.93	2.05	1.91
	IIB Context[b] (Å)	4.26	3.94	1.99
	IIA IIB Context[c] (Å)	3.56	3.07	1.55
Ecoli w. GrpE	IA IB Best[a] (Å)	1.32	1.39	1.15
	IIB Context[b] (Å)	3.42	3.41	4.20
	IIA IIB Context[c] (Å)	3.18	2.81	3.12

Footnotes:
a) RMSD of best CA super position of the secondary structure in IA and IB
b) RMS CA coordinate difference of the secondary structure in IIB, when IA and IB are best superposed
c) RMS CA coordinate difference of the secondary structure in IIA and IIB, when IA and IB are best superposed

Figures and Legends

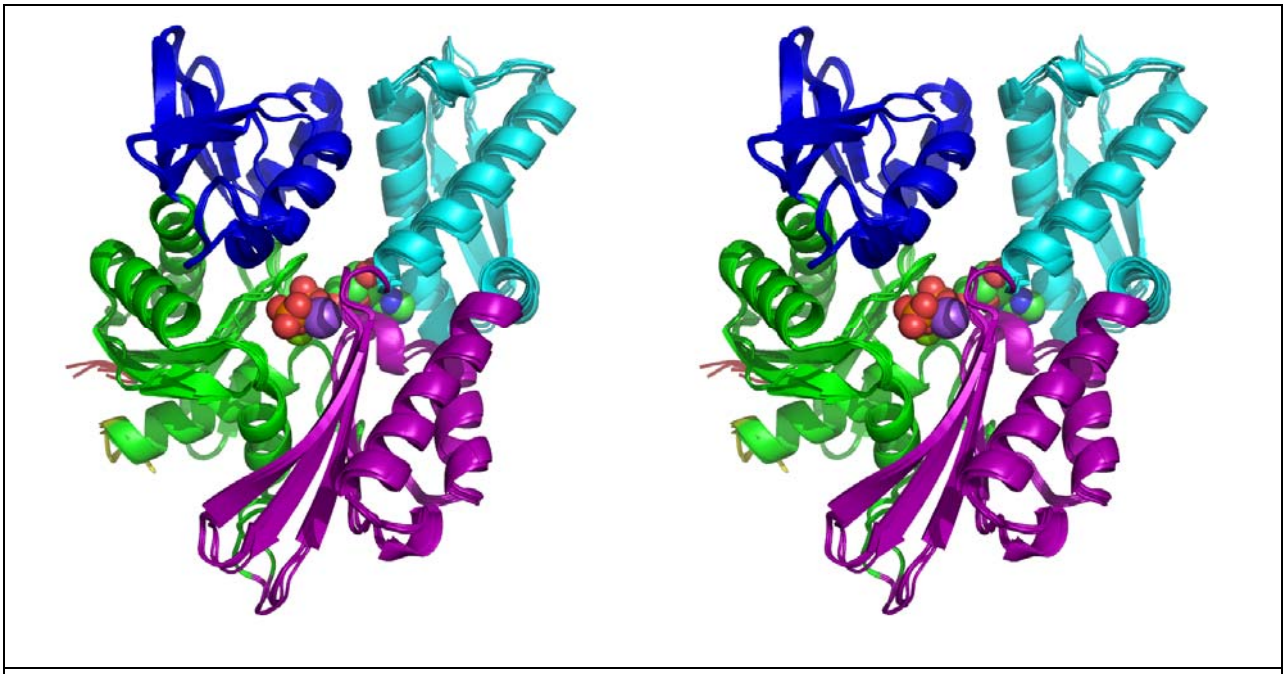

Figure 1:

Stereograph of the superposition of five X-ray crystallography structures for bovine Hsc70 NBD (4-380). The proteins were crystallized in the following states: wt-APO (2QW9.PDB, chain A), wt-ADP.PO_4 (3HSC.pdb and 2QWL.PDB, chain A), wt-ADP.VO_4 (2QWM.PDB, chain A) and K71M-ATP (1KAX.PDB). The proteins were overlaid on the secondary structure elements of sub domains IA (Hsc70-count residues 1-39,116-188, 361-381; green) and IB (residues 40-115; blue), to accentuate positional changes in sub domains IIA (residues 189-228, 307-360; purple) and IIB (residues 229-306; cyan). The N-terminus is in red, the C-terminus of this domain is shown in yellow. The nucleotide, PO_4^{3-}, Mg^{2+} and two Na^{+} present in 3HSC.pdb are shown in space fill.

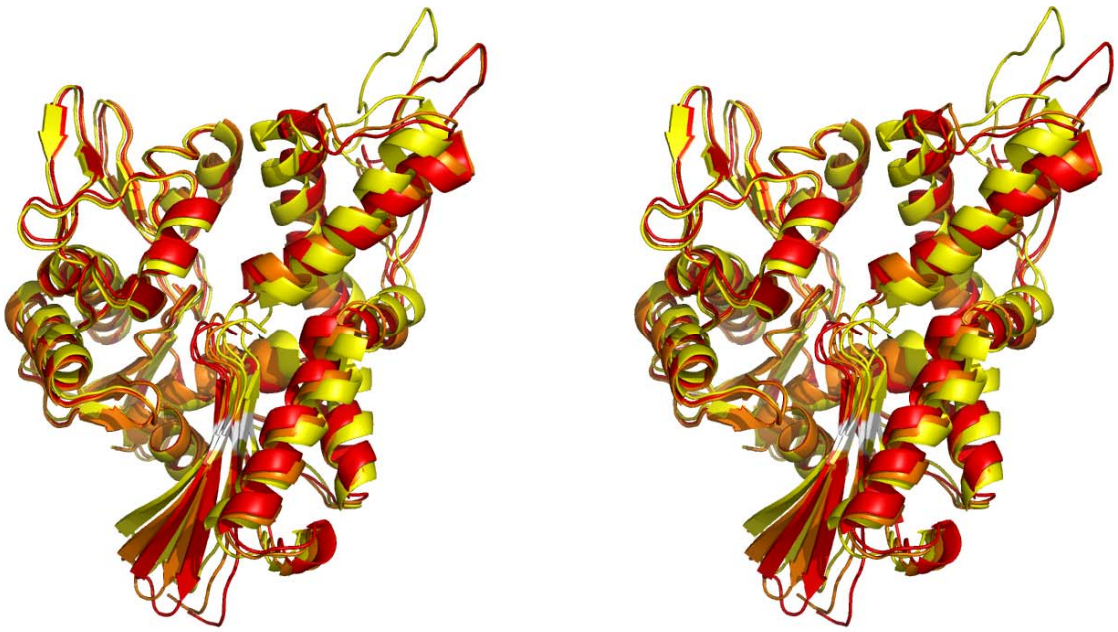

Figure 2.

Stereo graph of a superposition of the solution NMR (RDC) models for the NBD of DnaK-Tth, in the AMP-PNP state (yellow), the ADP.Pi state (orange) and APO state (red). The proteins were overlaid on the secondary structure elements of sub domains IA and IB (left half in this figure), to accentuate positional changes in sub domains IIA and IIB (right half in this figure).

White: putative node residues 189,198,214 and 333 (DnaK-Tth count) which form a hinge of a scissor-like motion.

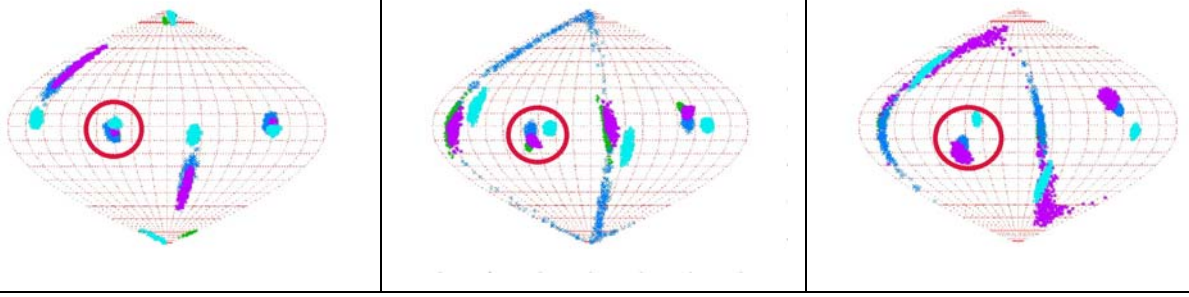

Figure 3:

Globe graphs (Sanson-Flamsteed plots) showing the orientations and experimental uncertainties of the S_{zz} principal alignment axes (in red circle around 60 ^{o}West and 5 ^{o}South) of the different subdomains of NBD of DnaK-Tth as derived from the NMR RDC measurements. Left, AMP-PNP state; middle, ADP.Pi state; right, APO state. Color code: sub domain IA (DnaK-Tth-count residues 1-37,109-180, 357-377; green); sub domain IB (residues 38-108; blue); sub domain IIA (residues 181-219, 307-356; purple) and sub domain IIB (residues 220-304; cyan). The reference structure for the NBD of DnaK-Tth was modeled on the NBD of bovine Hsc70.ADP.Pi (3HSC.PDB).

Scale: Horizontal: 20 ^{o} per gridline, vertical, 10 ^{o} per gridline. (The S_{zz} principal alignment axes also appear at the other side of the globe (120 ^{o}East, and 5 ^{o}North). The S_{xx} axes appear as smears around 150 ^{o} West, 45 ^{o} North and 30 ^{o}East, 45 ^{o}South. The polar presences are low abundant alternative 90 ^{o}rotated solutions)

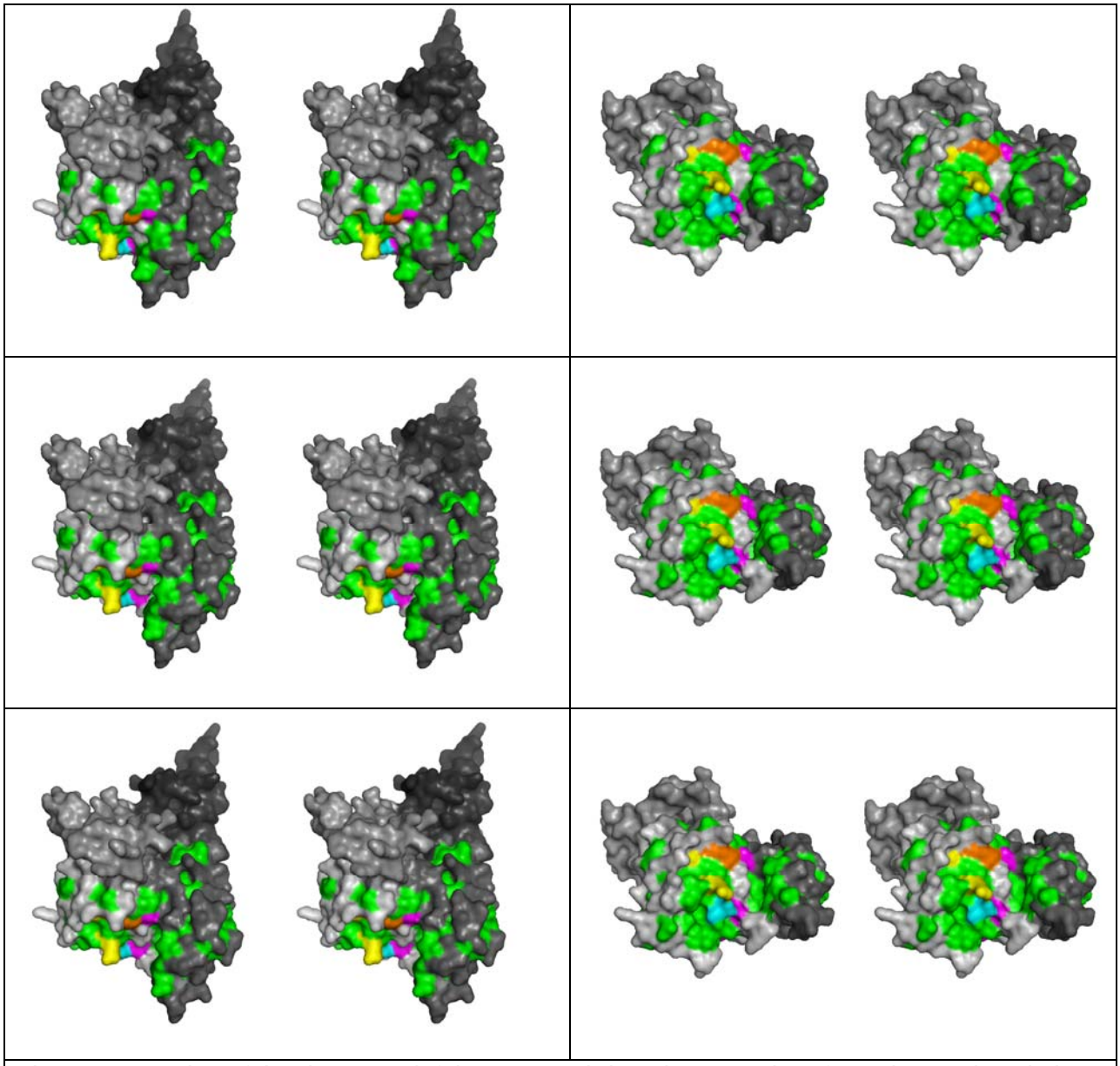

Figure 4. Opening of the “bottom” cleft between subdomains IA and IIA for TTh NBD in solution
Top row, AMP-PNP; middle row, ADP ; bottom row: APO , all in two orientations.
Color coding: Domains IA, IB, IIA, IIB are colored in the order light gray to dark gray.
Hydrophobics on IA and IIA are in green; C-terminus (residue 372) is in cyan;
Linker chemical shift sites[9] in magenta; mutation sites[32] 152+164 (TTH count) in yellow; pro switch site, residues 140+148, in orange[20].

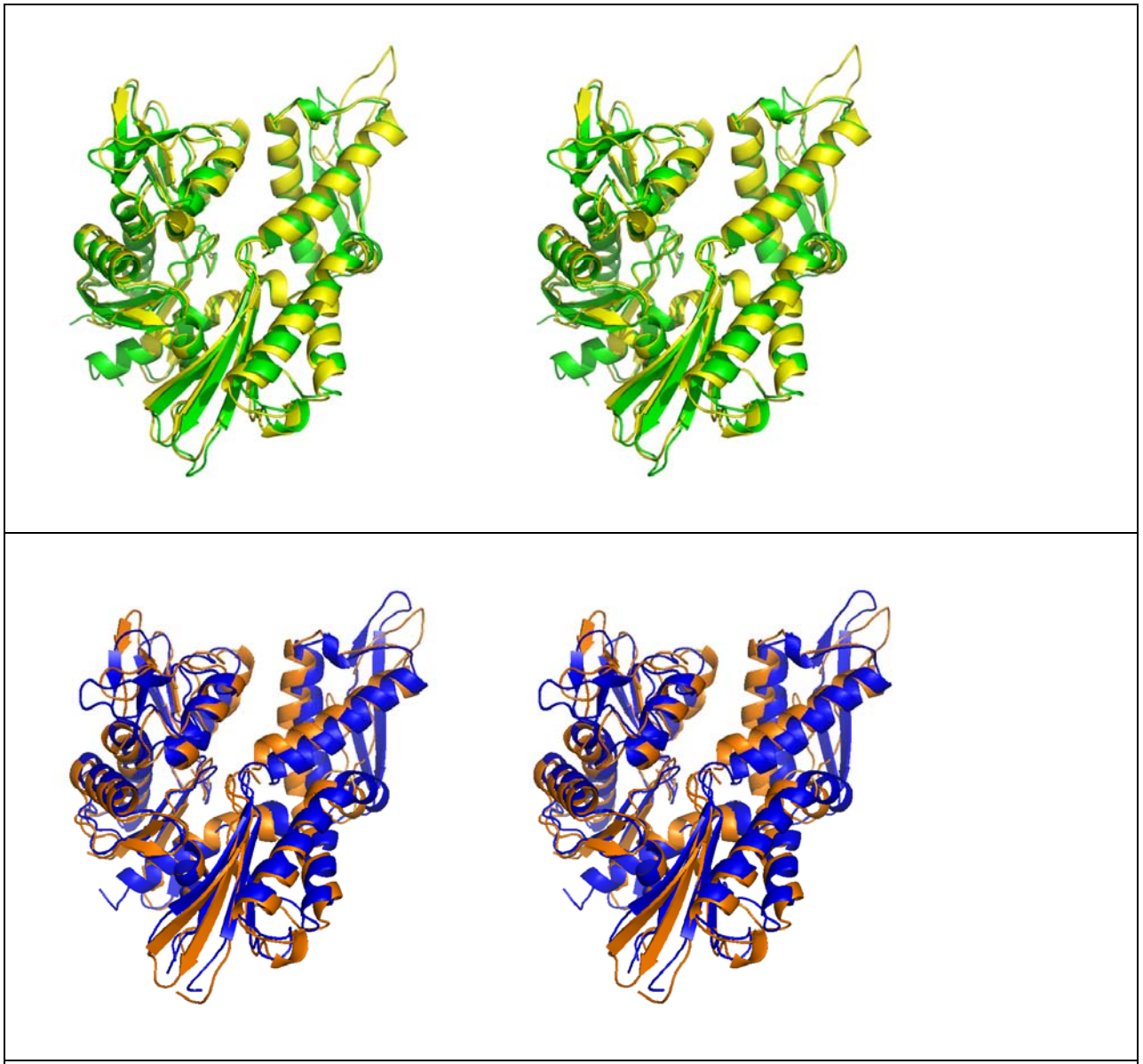

Figure 5.

Top: Superposition of solution Tth-DnaK-NBD-AMP-PNP (yellow) with crystal bovine Hsc70-NBD-ADP.Pi [4] (3HSC.pdb). The overlay of corresponding secondary structure elements of sub domains IA and IB (left) was optimized.

Bottom: overlay of solution Tth-DnaK-NBD-ADP (orange) with crystal E.coli-DnaK-NBD-APO from the complex with GrpE[23] in blue (GrpE not shown; 1DKG.pdb). The overlay of corresponding secondary structure elements of domains IA and IB (left) was optimized.