

## Human Inducible Hsp70: Structures, Dynamics, and Interdomain Communication from All-Atom Molecular Dynamics Simulations

Adrien Nicolaï,<sup>†</sup> Patrick Senet,<sup>\*,†</sup> Patrice Delarue,<sup>†</sup> and Daniel R. Ripoll<sup>‡,§</sup>

<sup>†</sup> Laboratoire Interdisciplinaire Carnot de Bourgogne, UMR 5209 CNRS-Université de Bourgogne, 9 Av. A. Savary, BP 47 870, F-21078 Dijon Cedex, France

<sup>‡</sup> Computational Biology Service Unit, Cornell Theory Center, Cornell University, Ithaca, New York 14853

Received April 23, 2010

**Abstract:** The 70 kDa human heat shock protein is a major molecular chaperone involved in de novo folding of proteins *in vivo* and refolding of proteins under stress conditions. Hsp70 is related to several “misfolding diseases” and other major pathologies, such as cancer, and is a target for new therapies. Hsp70 is comprised of two main domains: an N-terminal nucleotide binding domain (NBD) and a C-terminal substrate protein binding domain (SBD). The chaperone function of Hsp70 is based on an allosteric mechanism. Binding of ATP in NBD decreases the affinity of the substrate for SBD, and hydrolysis of ATP is promoted by binding of polypeptide segments in the SBD. No complete structure of human Hsp70 is known. Here, we report two models of human Hsp70, constructed by homology with *Saccharomyces cerevisiae* cochaperone protein Hsp110 (open model) and with *Escherichia coli* 70 kDa DnaK (closed model) and relaxed for several tens to hundreds of nanoseconds by using all-atom molecular dynamics simulations in explicit solvent. We obtain two stable states, Hsp70 with SBD open and SBD closed, which agree with experimental and structural information for ATP-Hsp70 and ADP-Hsp70, respectively. The dynamics of the transition from the open to closed states is investigated with a coarse-grained model and normal-mode analysis. The results show that the conformational change between the two states can be represented by a relatively small number of collective modes which involved major conformational changes in the two domains. These modes provide a mechanistic representation of the communication between NBD and SBD and allow us to identify subdomains and residues that appear to have a critical role in the conformational change mechanism that guides the chaperoning cycle of Hsp70.

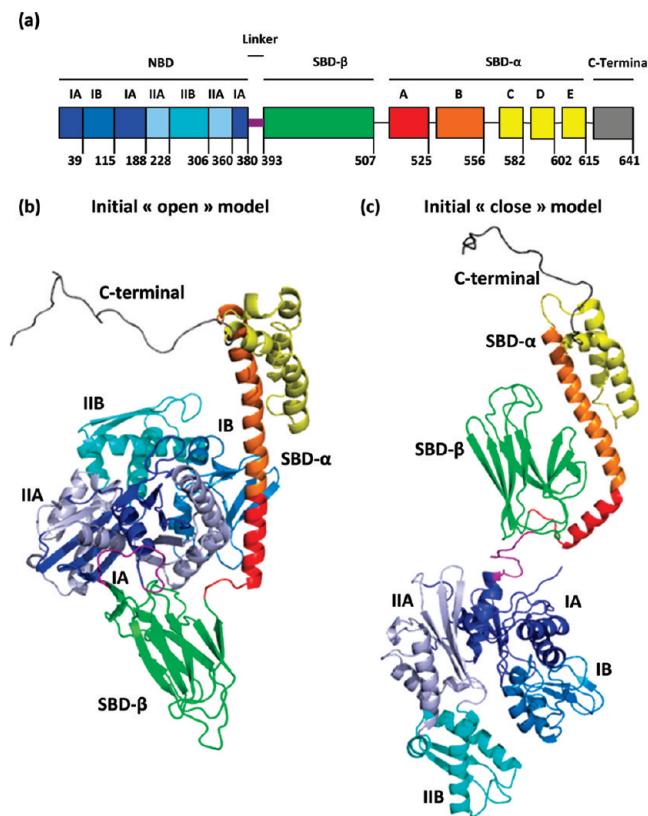
### 1. Introduction

The folding of many proteins *in vivo* is assisted by molecular chaperones.<sup>1–4</sup> The chaperones increase the efficiency of protein folding and inhibit the interactions of nascent proteins

extruded from the ribosome with other proteins within the intracellular medium.<sup>1–4</sup> The 70 kDa heat shock protein (Hsp70) is one of the major molecular chaperones involved in de novo folding of proteins *in vivo*. Hsp70 molecular chaperones are ubiquitous and found in eukaryotes and prokaryotes.<sup>1–4</sup> Hsp70 is overexpressed under cellular stress conditions, such as, for example, heat shock and pathologies,<sup>1–8</sup> and this molecular chaperone has been found associated with pathological proteins in several “protein misfolding diseases” in humans.<sup>9</sup> An increase of the expression of Hsp70 in

\* Corresponding author phone: +33 3 80 39 59 22; fax: +33 3 80 39 60 24; e-mail: psenet@u-bourgogne.fr.

<sup>‡</sup> Present address: Biotechnology HPC Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, 2405 Whittier Drive, Suite 200, Frederick, Maryland 21702



**Figure 1.** (a) Definition of the secondary structures of hHsp70. The color code is the following for the NBD (residues 1–380): subdomain IA (residues 1–39, 116–188, and 361–380; blue), subdomain IB (residues 40–115; marine), subdomain IIA (residues 189–228 and 307–360; light-blue), and subdomain IIB (residues 229–306; cyan). The SBD (residues 394–615) is formed by a  $\beta$  sandwich (SBD- $\beta$ ) and a helix bundle (SBD- $\alpha$ ). The SBD- $\alpha$  has two helices, A (residues 512–525; red) and B (residues 526–556; orange), which are in contact with the SBD- $\beta$  (residues 394–507; green) in the experimental structure of DnaK+peptide (PDB ID: 2KHO) and three other helices, C (residues 563–582; yellow), D (residues 589–602; yellow), and E (residues 605–615; yellow), forming a so-called hydrophobic core. The SBD is connected to the NBD by a short peptide (residues 381–393; purple) referred to as a linker. The C-terminal part (residues 616–641; gray) is a quite unstructured chain of amino acids terminated by the EEVD chaperone motif (residues 638–641; gray). (b) Ribbon diagram of the initial open model built by homology from the Hsp110 structure (PDB ID: 3C7N chain A). The color code is similar to the one in panel a. (c) Ribbon diagram of the initial closed model built by homology from the DnaK structure (PDB ID: 2KHO). The color code is similar to the one in panel a. This figure was prepared with PyMOL [<http://www.pymol.org>].

models of misfolding diseases has been shown to result in a decrease of the toxicity of disease proteins and/or to inhibit disease protein aggregation.<sup>9–13</sup> In addition, Hsp70 interacts with hundreds of client proteins of the intracellular environment.<sup>1,5,8,14,15</sup>

Hsp70 proteins are comprised of a nucleotide binding domain (NBD) and a substrate binding domain (SBD). The NBD is divided into two rather symmetrical lobes, I and II, each divided into two subdomains, A and B (Figure 1a). The SBD has a peptide binding pocket (this subdomain is referred

to as SBD- $\beta$ ) and an  $\alpha$ -helical “lid” (this subdomain is referred to as SBD- $\beta$ ; Figure 1a) which is believed to control the access of the substrate to the binding cavity.<sup>1–5</sup> The position of SBD- $\alpha$  relative to SBD- $\beta$  defines two main conformational states of Hsp70 chaperone:<sup>1–5</sup> either the lid is open and the peptide can access the hydrophobic pocket within SBD- $\beta$ , named the “open” conformation of the chaperone (as in the initial model shown in Figure 1b), or the lid is closed and the peptide is trapped in the pocket, referred to as the “closed” conformation (as in the initial model shown in Figure 1c).

Hsp70 assists the folding of other proteins, through cycles of binding and release of unfolded polypeptide chains in the SBD by binding selectively short peptide stretches within the chain.<sup>1–3</sup> The binding and release of peptides is governed by ATP hydrolysis in the NBD and by nucleotide exchange<sup>16,17</sup> and is promoted by the action of cochaperone proteins.<sup>18–20</sup> The communication between the SBD and NBD, which regulates the binding and release of peptides, is governed by a linker (Figure 1a) through a mechanism which is not fully understood.<sup>21,22</sup> In ATP-Hsp70, the affinity for a substrate is low,<sup>21</sup> and therefore the molecular chaperone is expected to be in an open state (as in the initial model shown in Figure 1b). On the contrary, in ADP-Hsp70, the affinity for a substrate is high,<sup>21</sup> and the chaperone is expected to be thus in a closed state (as in the initial model shown in Figure 1c).

Because of their critical roles in many cellular processes and based on evidence that their mechanisms involve cycles of conformational changes, there has been strong interest in elucidating the structures of Hsp70's. Crystal structures from bovine and human NBDs and NMR structures and crystal structures of SBDs from rat Hsc70 and *E. coli* Hsp70 (DnaK) have been solved during the past few years.<sup>23–30</sup> More recently, structures comprising both the NBD and SBD have become available; a truncated structure of the bovine Hsc70 chaperone (residues 1–554) without a nucleotide and substrate<sup>31</sup> and a truncated structure of *E. coli* DnaK (residues 4–603), complexed with ADP nucleotide and a short peptide (NRLLLTG), have been reported.<sup>32</sup> Furthermore, a crystal structure of *Saccharomyces cerevisiae* Hsp110 (residues 2–659), which is an Hsp70 homologue, was recently elucidated.<sup>20</sup> In this structure, Hsp110 was crystallized in an ATP state with the SBD open. Hsp110 is a cochaperon of Hsp70. It is a nucleotide-exchange factor (NEF) which binds to the NBD of the target chaperons<sup>18–20</sup> and induces ADP release from the NBD.

There is no structure of an Hsp70 locked in the ATP state (except the yeast Hsp110 homologue<sup>20</sup>). In addition, there are no structural data reporting the structure of the SBD of human Hsp70. Consequently, the alternation of conformations between the ATP and ADP states of a full length (human) Hsp70 has not been described at the atomic level. Prompted by the absence of a detailed atomic experimental structure, we carried out the task of generating computer models of human Hsp70 (hHsp70) in two conformational states: with the lid open (Figure 1b) and with the lid closed (Figure 1c). The sequence of human inducible Hsp70 (UniProt ID: P08107) was chosen for modeling purposes

because it has been reported as a new target for cancer therapies.<sup>15,33–35</sup>

Here, we first report a realistic atomic model of the complete structure of the inducible Hsp70 (residues 1–641) in an open conformation that was built from the yeast Hsp110 structure<sup>18</sup> (PDB ID: 3C7NA) using homology modeling (Figure 1b) and relaxed through several hundred nanoseconds by using all-atom molecular dynamics (MD) simulations in explicit solvent. These simulations resulted in a structural model for the open conformation of Hsp70, which agrees with the ATP state of hHsp70 when compared to the low-resolution small-angle X-ray scattering (SAXS) data.<sup>36,37</sup> We next report a second model of the human Hsp70 chaperone in a closed conformation that was built using homology modeling (Figure 1c) from the experimental structure of *E. coli* DnaK<sup>32</sup> (PDB ID: 2KHO). This second model was relaxed through hundreds of nanoseconds of all-atom MD and converged to a stable structural model of hHsp70. The latter model agrees with the ADP state of hHsp70 when compared to the low-resolution SAXS data.<sup>36,37</sup>

The transition between the open and closed conformations of hHsp70 was analyzed by computing the collective vibrational modes<sup>38,39</sup> of the two atomic models obtained by MD representatives of the open and closed conformations, respectively, using a coarse-grained model.<sup>40–46</sup> We show conclusively that a few collective modes of Hsp70 contribute to the interpolated conformational “pathway” from ATP-Hsp70 to ADP-Hsp70. In addition, these collective modes provide a possible mechanism to propagate information between the NBD and the SBD at a distance as large as 100 Å.

## 2. Methods

### 2.1. Generation of the Initial 3D Models of Hsp70.

Three dimensional models of the human Hsp70 protein were generated by using homology modeling techniques. A BLAST<sup>47</sup> search using the sequence of Hsp70 against all sequences in the Protein Data Bank [i.e., sequences for which three-dimensional (3D) information has been collected] identifies few highly homologous sequences. Among these possible templates with the largest degree of residue identity (RI) detected by BLAST were 3C7N monomer B, 1YUW monomer A, and 2VZ7 monomer A. However, the actual 3D data from these structural determinations do not fully cover the whole sequence range of our protein of interest [e.g., for 3C7N\_B from a total of 641 residues in Hsp70, only 554 are aligned (~86%)]. On the other hand, BLAST was able to detect few other sequences that, while sharing lower RI, provided better coverage (i.e., a larger number of aligned residues). Among these structures are (a) 2KHO monomer A, for which the alignment provides coverage for 584 of 641 residues in the Hsp70 sequence (91%), with RI = 310/611 (50%), 402/611 (65%) residues with positive scores, and 26/611 (4%) gaps; and (b) 3C7N monomer A, for which the alignment provides coverage for 598 residues out of 641 (or 93%), with RI = 183/646 (28%), 315/646 (48%) residues with positive scores, and 49/646 (7%) gaps. On the basis of these results, we decided to use the

experimentally determined structure of the *E. coli* Hsp70 (DnaK) chaperone<sup>32</sup> (PDB ID: 2KHO monomer A) and the structure of the heat shock protein Hsp110 homologue Sse1<sup>18</sup> (PDB ID: 3C7N monomer A) as templates to produce a set of 3D models for the human Hsp70 sequence.

To carry out this task, we made use of the program MODELLER<sup>48–50</sup> (version 9.3), which is capable of converting the information from (a) pairwise alignments of the query sequence with its templates and (b) the structural data provided by these templates into a set of distance and dihedral-angle restraints. By minimizing the violations to these derived restraints, MODELLER<sup>48–50</sup> is able to generate the 3D models for the sequence of interest. Initial models were energy minimized using 500 to 1000 steps of a steepest-descent algorithm provided in the DS modeling program (Accelrys, Inc.). The optimized models showing the closest C<sup>α</sup> RMSD (root-mean-square deviation) with respect to the templates from which they were derived were used as starting points for MD simulations. In particular, the structural superposition of the open model of the APO-Hsp70 used as an initial conformation of one of the MD runs (APO-1 run), and its template, the structure of Hsp110, shows very small differences with a C<sup>α</sup> RMSD of 0.8 Å for an optimal superposition of 609 C<sup>α</sup> atoms. Similarly, the structural superposition of the MD initial conformation for the closed model of Hsp70 (APO run) and its template, the structure of DnaK, led to a C<sup>α</sup> RMSD of 1.6 Å for an optimal superposition of 590 C<sup>α</sup> atoms. Structure alignments were carried out by using the combinatorial extension method.<sup>51</sup>

**2.2. Molecular Dynamics Simulations.** All-atom molecular dynamics simulations in explicit water of Hsp70 have been carried out with the GROMACS software package<sup>52</sup> using the simple point charge (SPC) water model and the GROMOS96 ffG43a1 force field.<sup>52</sup> The time step used in all simulations was 0.001 ps, and the list of neighbors was updated every 0.005 ps with the “grid” method and a cutoff radius of 1 nm. The coordinates of all the atoms in the simulation box were saved every 2 ps. The initial velocities were chosen randomly. We used the NPT ensemble with a cubic box. The temperature and pressure were kept to the desired value by using the Berendsen method and an isotropic coupling for the pressure ( $T = 300$  K,  $\tau_T = 0.1$  ps;  $P_0 = 1$  bar, coupling time  $\tau_P = 1$  ps). The electrostatic term was computed by using the particle mesh Ewald (PME) algorithm (with a radius of 1 nm) with the fast Fourier transform optimization (on 117 points for each axis for the open model and 160 points for the closed model and an order equal to 4 for the interpolation). The cutoff algorithm was applied for the noncoulomb potentials with a radius of 1 nm. The system was warmed up for 40 ps and equilibrated for 600 ps with lower restraints, finishing with no restraints at 300 K. We performed five runs, four for the open model and one for the closed model.

**Open Model. APO-1 Run.** The initial structure of APO-Hsp70 (open model, run APO-1) was built by using the structure of Hsp110 as a template.<sup>18</sup> The APO-Hsp70 protein has been solvated in a cubic box with 82 065 SPC water molecules keeping a minimum distance of 0.9 nm between the solute and each face of the box. We used the periodic

boundary conditions, and the initial value of the box side is 13.683 nm. The charge of APO-Hsp70 has been neutralized by adding 11 Na<sup>+</sup> counterions. The energy of the model was first optimized with the “steepest descent minimization” algorithm and then by using the “conjugate gradients” algorithm. The production period was 400 ns.

**ADP-1 Run.** The ADP-Hsp70 model was built using a structure of Hsp70 relaxed after a 3 ns period of production of run APO-1. We added one ADP molecule with PRO-DRG<sup>53</sup> and one Ca<sup>2+</sup> ion (as in the experimental structure of the human Hsp70 NBD, PDB ID: 1HJO<sup>26</sup>) to this APO structure. The protein was solvated with 82 089 water molecules in a cubic box of the same dimensions as in the APO-1 run, and 12 Na<sup>+</sup> ions were added to neutralize the system. The production period was 480 ns.

**APO-2 Run.** A second initial structure of APO-Hsp70 was built as follows. We observed that the initial conformation of ADP-Hsp70 was stable until 320 ns. The “A+B” helix did not break into two helices before 320 ns. To test the effect of the nucleotide and to produce different initial conditions for the APO form, we selected the structure of ADP-Hsp70 at 200 ns (before any conformational change). We removed the ADP molecule, the Ca<sup>2+</sup> ion, and one Na<sup>+</sup> ion, and we added eight water molecules (with initial velocities equal to zero; the velocities of all the other atoms in the ADP-Hsp70 structure at 200 ns were conserved). This new APO structure was the initial structure of the APO-2 run. The production period was 184 ns.

**ADP-2 Run.** A second initial structure of ADP-Hsp70 was built similarly to the first trajectory ADP-1 but with different initial velocities. The production period was 192 ns.

**Closed Model. APO Run.** The initial structure of Hsp70 (closed model, APO run) was built by using the structure of DnaK as a template.<sup>32</sup> The closed model Hsp70 protein was solvated in a cubic box with 212 632 SPC water molecules keeping a minimum distance of 0.9 nm between the solute and each face of the box. We used periodic boundary conditions with an initial value of the box side of 18.664 nm. The charge of the system in the closed model was neutralized by adding 11 Na<sup>+</sup> counterions. The energy of the system was first optimized with the “steepest descent minimization” algorithm and then by using a “conjugate gradient” algorithm. The production period was 130 ns.

**2.3. The Anisotropic Network Model (ANM).** An important issue in the analysis of protein dynamics is the identification of the mechanism of slow, large-amplitude motions, also termed collective motions. The collective motions are believed to be essential for a protein to function. A widely used technique for studying collective motions of large proteins is the anisotropic network model<sup>41–46</sup> (ANM). In ANM, a protein in its folded state is represented by a three-dimensional elastic network of nodes, which are located at the C<sup>α</sup> position of each residue. The interactions between residues in close proximity (distance < cutoff distance  $r_c$ ) are replaced by harmonic springs.<sup>41</sup> The ANM predicts the relative sizes of structural fluctuations of a protein in its native state and their directionalities and permits the decomposition of the molecular motions into a series of  $3N - 6$  modes ( $N = 641$  for the human Hsp70).

In ANM, the intramolecular free energy of a protein of  $N$  residues in its native state is expressed as a series expansion in the fluctuations  $\Delta\vec{R}_i$  of individual residue positions ( $1 \leq i \leq N$ )

$$F - F(0) = \sum_{i=1}^N \left( \frac{\partial F}{\partial \Delta\vec{R}_i} \right)_0 \cdot \Delta\vec{R}_i + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left( \frac{\partial^2 F}{\partial \Delta\vec{R}_i \partial \Delta\vec{R}_j} \right)_0 : \Delta\vec{R}_i \Delta\vec{R}_j = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \Phi_{ij} \cdot \Delta\vec{R}_i \Delta\vec{R}_j \quad (1)$$

where the first term on the right-hand side of the first line equality is zero because, in the native state

$$\left( \frac{\partial F}{\partial \Delta\vec{R}_i} \right)_0 = 0, \forall i$$

The  $3N \times 3N$   $\Phi$  matrix in eq 1 is defined as follows

$$\Phi_{ij}^{\alpha\beta} = \left( \frac{\partial^2 F}{\partial \Delta\vec{R}_i \partial \Delta\vec{R}_j} \right) = -A \frac{(R_i^\alpha - R_j^\alpha)(R_i^\beta - R_j^\beta)}{R_{ij}^2} \Gamma_{ij} \quad (2)$$

where  $R_i^\alpha$  is the Cartesian component  $\alpha$  ( $= x, y$ , or  $z$ ) of the position  $\vec{R}_i$  of the  $i$ th C<sup>α</sup> atom and  $A$  and  $\Gamma_{ij}$  are parameters of the ANM model, which depend on the protein studied. Equation 2 is the simplest form of  $\Phi$ , which is invariant by global translation and rotation of the molecule.

The elements  $\Gamma_{ij}$  of the matrix in eq 2 are dimensionless and take the value 1 if the inter-residue distance  $R_{ij}$  remains below the cutoff distance  $r_c$  and is 0 otherwise. The parameter  $A$  is the average force constant between two connected residues. The cross-correlations between residue fluctuations are found from

$$\langle \Delta R_i^\alpha \cdot \Delta R_j^\beta \rangle = kT[\Phi^{-1}]_{ij}^{\alpha\beta} \quad (3)$$

where  $[\Phi^{-1}]_{ij}^{\alpha\beta}$  is the  $\alpha\beta$  element of the  $ij$ th  $3 \times 3$  bloc matrix of the  $3N \times 3N$  inverse matrix  $\Phi^{-1}$ .

The mean square fluctuation of the bonds  $R_{ij}$  is computed from eq 3, with  $i = j$ , i.e.,

$$\langle (\Delta R_i)^2 \rangle = kT \sum_{\alpha} [\Phi^{-1}]_{ii}^{\alpha\alpha} \quad (4)$$

Equation 3 will be used here to calculate the  $B$  factors ( $B$ ) of each residue:

$$B_i = \frac{8\pi^2}{3} \langle (\Delta R_i)^2 \rangle \quad (5)$$

The matrix  $\Phi$  is rewritten as the product of three matrices, the diagonal matrix  $\Lambda$  of its nonzero eigenvalues  $\lambda_m$  ( $1 \leq m \leq 3N - 6$ ), the matrix  $\mathbf{E}$  of the corresponding eigenvectors  $\vec{e}_m$ , and the transpose of  $\mathbf{E}$ :

$$\Phi = \mathbf{E} \Lambda \mathbf{E}^{-1} =$$

$$[e_1 e_2 \dots e_{3N-6}] \text{ diag}(\lambda_1 \lambda_2 \dots \lambda_{3N-6}) [e_1 e_2 \dots e_{3N-6}]^T \quad (6)$$

where  $\mathbf{E}$  is an orthonormal matrix. The inverse of  $\Phi$  can be readily found from eq 6 and is given by

$$\Phi^{-1} = \mathbf{E}\Lambda^{-1}\mathbf{E}^{-1} = \sum_m \left[ \frac{\vec{e}_m \cdot \vec{e}_m^T}{\lambda_m} \right] \quad (7)$$

The eigenvalues of  $\Phi$  are ordered as  $\lambda_1 < \lambda_2 < \lambda_3 \dots < \lambda_n$  where  $\lambda_1$  is the first nonzero eigenvalue. These eigenvalues  $\lambda$  have the physical dimension of a restoring force but are also called frequencies by analogy with the normal modes analysis. The collective modes correspond to low  $\lambda$ . The number of modes of nonzero frequency for hHsp70 is 1917, and the slow modes here are defined as the first hundred modes.

The covariance of the structural fluctuations for the pair of residues  $i-j$  is related to the spectrum of  $\Phi$  by

$$\langle \overrightarrow{\Delta R_i} \cdot \overrightarrow{\Delta R_j} \rangle = kT \sum_m \left[ \frac{\vec{e}_m \cdot \vec{e}_m^T}{\lambda_m} \right]_{ij} \quad (8)$$

The quantity  $1/\lambda_m$  is the reciprocal of the  $m$ th nonzero eigenvalue  $\lambda_m$  of  $\Phi$ . According to eqs 4 and 8, the slowest mode (with the lowest frequency  $\lambda_1$ ) has the most dominant contribution to the structural fluctuations of the protein and corresponds to displacements along  $\vec{e}_1$ . The slowest modes are thermally excited with a large amplitude and generally describe functional motions that are of great biological interest.

The ANM requires specification of two parameters: the force constant  $A$  that is a measure of the strength of the restoring forces of the  $C^\alpha-C^\alpha$  virtual bonds of the native fold (eq 2) and the cutoff distance  $r_c$  that defines whether two given nodes are linked or not. The values of  $A$  and  $r_c$  in ANM were fitted in order to reproduce the  $B$  factors extracted from the crystallographic data of the truncated bovine Hsc70<sup>31</sup> (bHsc70) (PDB ID: 1YUW; data not shown). The value of the cutoff distance  $r_c$  was obtained as follows. We noted that there were more than six zero eigenvalues in the ANM model for  $r_c \leq 10 \text{ \AA}$  for both the bHsc70 structure and for the model average structures (computed between  $t = 100 \text{ ns}$  and  $t = 400 \text{ ns}$  for the open model and computed between  $t = 30 \text{ ns}$  and  $t = 130 \text{ ns}$  for the closed model). This implied that some deformation of the structure did not cost internal energy: for these deformations, the structure was unstable. To remove such physically unrealistic behavior, we adopted a larger cutoff distance, which was evaluated by studying the variation of the spectrum of  $\Phi$  for the bHsc70 structure for different values of  $r_c$ . In particular, we computed the density of states,  $g(\lambda)$ , of the low frequency modes. The distribution  $g(\lambda)$  was found to increase linearly with the frequency<sup>54</sup> for the slow modes (defined here as the first 100 modes which contribute to about 70% of the total fluctuations), implying that the cumulative density of modes,  $G(\lambda) = \int_0^\lambda du g(u) \propto \lambda^2$ , deviated from the Debye model of elastic solids where  $G(\lambda) \propto \lambda^3$ . The anomalous spectral dimension  $d_s \approx 2$  of proteins, defined as the power of  $\lambda$ , i.e.,  $G(\lambda) \propto \lambda^{d_s}$ , was observed by inelastic neutron scattering for lysozyme,<sup>55</sup> for which  $d_s \approx 1.4$ . The spectral dimension of bHsc70 for each  $r_c$  value between 11 and 20  $\text{\AA}$  was computed from a power law that best fits the cumulative density of modes for the first 100 modes. The results showed that the

value of  $d_s$  is nearly constant ( $\approx 1.9$ ), and the shape of the density of modes  $g(\lambda)$  does not change for  $11 \leq r_c \leq 14 \text{ \AA}$ . Therefore, we have chosen a  $r_c = 11 \text{ \AA}$ . The best value of  $A$  reproducing the  $B$  factors of bHsc70<sup>31</sup> for  $r_c = 11 \text{ \AA}$  was 0.4–0.5 kcal/mol/ $\text{\AA}^2$ .

Another parametrization of ANM was tested on the basis of the so-called parameter-free ANM<sup>45</sup> (pfANM), published during the course of the present work. In pfANM, there is no cutoff distance  $r_c$ , and the force constant  $A$  decays as an inverse power  $n$  of the distance between a pair of  $C^\alpha$ 's. A value of  $n \geq 6$  is recommended to reproduce the collective modes of a protein. The collective modes computed with pfANM were compared to the collective modes computed with ANM and  $r_c = 11 \text{ \AA}$  (data not shown). The collective modes were very similar to each other. The overlap of the eigenvectors of the first 10 nonzero frequency modes was greater than 0.93, both for the open and for the closed models of Hsp70 (data not shown).

**2.4. Involvement Coefficient.** In order to evaluate the contribution of a given mode to the transition between the open and closed states, individual and cumulative involvement coefficients<sup>56</sup> were calculated as follows.

A transition pathway was determined by linearly interpolating the transition end-points (i.e., the initial,  $\vec{X}_A$ , and final,  $\vec{X}_B$ , structures of a conformational transition) upon optimal superposition of all the  $C^\alpha$  atoms. The resulting “displacement vector”  $\vec{X}_A - \vec{X}_B$  was expanded in a linear combination of normal modes  $\vec{e}_k$  of the initial state. Then, the projection  $L_k$  of the normalized displacement vector on the  $k$ th normal mode vector was computed:

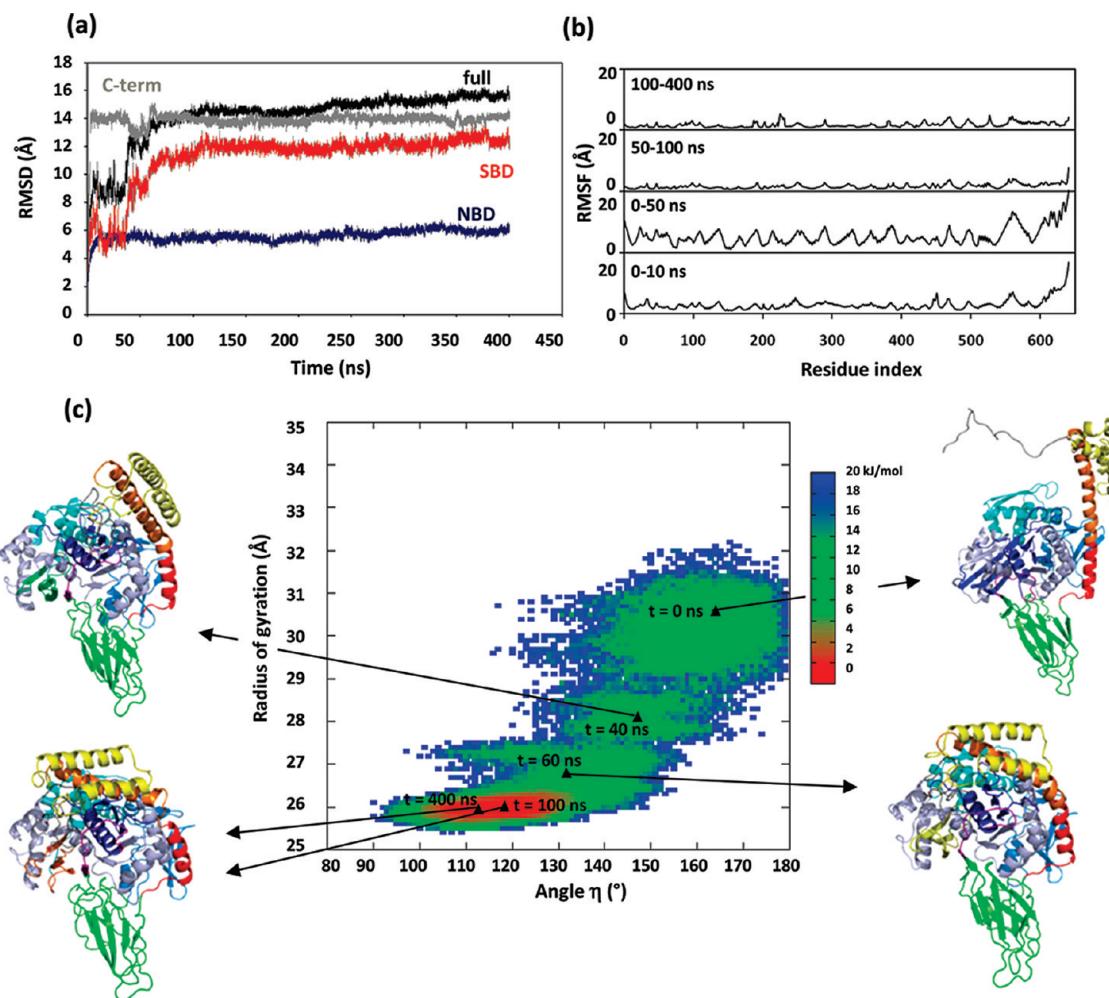
$$I_k = \frac{\vec{X}_A - \vec{X}_B}{|\vec{X}_A - \vec{X}_B|} \cdot \vec{e}_k \quad (9)$$

The involvement coefficient is defined as  $I_k = |I_k|$ . It describes the degree of involvement of the  $k$ th mode in the conformational transition  $A \rightarrow B$ . Thus, the individual involvement coefficient indicates in a semiquantitative way which collective motions are important for a given conformational change. A complementary quantity which indicates the involvement of a set of collective modes in the conformational change is the cumulative involvement coefficient  $CI_k$ , which is computed as

$$CI_k = \sum_{k=1}^N I_k^2 \quad (10)$$

### 3. Results and Discussion

**3.1. Structures of hHsp70 Relaxed by MD in an Open State. Initial Model Built by Homology.** The initial model of the human inducible Hsp70 structure in an open state (named the open model hereafter) was built by homology using the experimental structure of Hsp110<sup>18</sup> as a template in the software MODELLER<sup>48–50</sup> (see Methods). There is a main relevant structural difference between Hsp110 and all of the known Hsp70 truncated structures: in the SBD- $\alpha$  of Hsp110,<sup>18</sup> the helices A and B (Figure 1a) form a unique long helix (Figure 1b) that we call “A+B”.

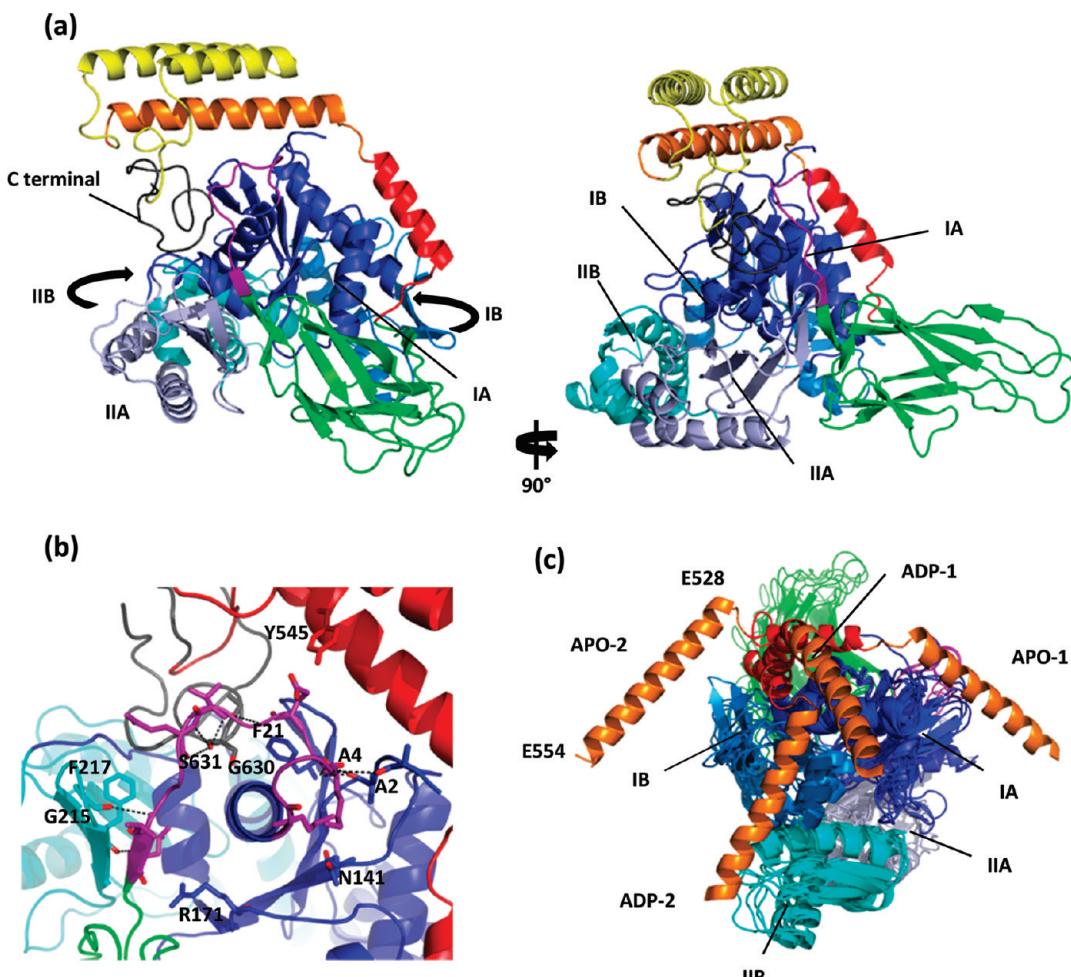


**Figure 2.** Time evolution of structural parameters of the open model as a function of time (run APO-1 is shown). (a) The  $C^\alpha$  RMSD (for the full structure, the NBD, the SBD, and the C-terminal part) computed with respect to the initial open model built by homology modeling. (b) The  $C^\alpha$  RMSF ( $\sqrt{\langle \Delta R_i^2 \rangle}$ , see eq 4 in Methods) computed for the open model in different time windows. (c) Energy surface (in kJ/mol) for the relaxation of the open model studied by MD. The energy surface is defined by  $E = -kT \ln[P(R_g, \eta)/\text{Max}\{P(R_g, \eta)\}]$  where  $P(R_g, \eta)$  is the probability density of the radii of gyration  $R_g$  and of the angle  $\eta$  defined in section 3.1 and computed from MD.  $E$  is zero at the maximum value of  $P(R_g, \eta)$  (red basin).

To the best of our knowledge, there is no structure of isolated Hsp110, and its long helix might be formed when Hsp110 interacts with its target protein<sup>18</sup> (Hsp70) or with itself, as shown in a crystal of the Hsp110 dimer<sup>20</sup> (PDB ID: 2QXL).

**MD of the Open Model without Nucleotide.** The initial model without nucleotide (APO-hHsp70) was relaxed by all-atom MD at 300 K (see Methods). The structural changes and the stability of the structure simulated by MD have been monitored by computing the  $C^\alpha$  RMSD with respect to the initial model (both for the full structure and for the SBD, NBD, and C-terminal domains separately) as a function of time for two trajectories of hHsp70 without a nucleotide (named APO-1 and APO-2, see Methods), as illustrated in Figure 2a for the APO-1 trajectory. The structure of the NBD simulated by MD converges to a stable conformation in less than 10 ns, as shown by the convergence of the  $C^\alpha$  RMSD with respect to the initial model in Figure 2a. The relatively huge RMSD value between the relaxed open model and the initial model found at 400 ns (16 Å; Figure 2a) is due to the reorganization of the SBD and of the C-terminal part and to the reorientation of the SBD relative to the NBD. Analysis

of the  $C^\alpha$  RMSF (root-mean-square fluctuations) along the primary sequence (Figure 2b) showed that the SBD structure and the C-terminal part of the initial open model built by homology were not stable. Indeed, we observed huge fluctuations of the SBD-α and of the C-terminal loops (residues index  $\geq 525$  in Figure 2b), reflecting a reorganization of the structure up to about 50 ns of simulation time in the MD trajectory APO-1 (Figure 2b). The relaxation of the C-terminal part of the initial model of hHsp70 occurred quickly (in few nanoseconds; Figure 2a) and induced unfolding of the helix E of its SBD-α (Figure 3a). The main structural changes that occurred at about 50 ns (shown as a sharp increase of the  $C^\alpha$  RMSD at 50 ns in Figure 2a for the SBD and for the full structure) were the break of the initial fused helix “A+B” of the SBD-α<sup>18,20</sup> into two helices A and B and the binding of the SBD-α on the NBD, as shown in Figure 3a. There was still a slow structural evolution of the model between 100 and 400 ns (as shown by a slow evolution of the  $C^\alpha$  RMSD of the full structure in Figure 2a), but the structures (not shown) computed between 100 and 400 ns were very similar.



**Figure 3.** Typical structures of the open model relaxed by MD simulations. The color code is similar to the one in Figure 1a. Ribbon diagram of the average open model (run APO-1) in two views rotated by 90°. (b) Close-up view of the contacts between the linker and the other domains. (c) Comparison of the average structures of the open model relaxed by MD: run APO-1, run APO-2, run ADP-1, and run ADP-2. The structures are truncated after residue 554 to show clearly the position of the SBD- $\alpha$ . The superposition of the structures was done by minimizing the RMSD of the  $C^\alpha$  of the NBD. This figure was prepared with PyMOL [<http://www.pymol.org>].

The reorganization of the structure of the initial open model was best represented by the energy surface  $E(R_g, \eta)$ ; Figure 2c) defined as an effective potential of mean field:  $E(R_g, \eta) \equiv -kT \ln[P(R_g, \eta)/\text{Max}\{P(R_g, \eta)\}]$ , where  $k$  is the Boltzmann constant,  $T$  is the temperature, and  $P(R_g, \eta)$  is the 2-D residential probability density<sup>57</sup> of the radius of gyration of the protein ( $R_g$ ) and of the angle measuring the local curvature of the “A+B” helix ( $\eta$ ) in the region where the helix was found bent in the relaxed structure. The angle  $\eta$  was calculated as the angle between the vector joining the  $C^\alpha$  of the residues 512 and the  $C^\alpha$  of the residue 516 at the N-terminal part of the “A+B” helix and the vector joining the  $C^\alpha$  of the residues 530 and the  $C^\alpha$  of the residue 534 of the helix B. The four-residue spacing used to define the  $C^\alpha-C^\alpha$  vectors was chosen because in a perfect  $\alpha$  helix a rotation of over 360° corresponds to a 3.6 residue spacing; in this way, the two vectors chosen are almost parallel to the helical axis with an angle of almost 180° in the initial model. The sequence of the major events of the reorganization of the model studied by MD is shown in Figure 2c: the initial break of the helix occurred at about 40 ns, and between 40 and 60 ns, the helix curvature increased (the angle  $\eta$

decreased) while simultaneously the SBD- $\alpha$  rotated, implying a stronger interaction between the SBD and the NBD. Binding of the helix on the NBD is observed around 100 ns. The structure found at 100 ns is closely similar to the one at 400 ns (the  $C^\alpha$  RMSD between the structures at 100 and 400 ns is 3.2 Å). The energy scale computed from canonical MD simulations, such as the one shown in Figure 2c, is approximate because the conformational sampling of canonical MD simulations is necessarily limited for a system of the size of Hsp70, and because the energy barriers in a 2-D energy surface may not represent the actual energy barriers in the multidimensional conformational space of the protein.

In the stable average structure computed between 100 and 400 ns, the helices A and B interact with the lobe IA of the NBD (Figure 3a). The interdomain linker (residues 381–393) is bound to a cleft between subdomains IA and IIA of the NBD (Figure 3b). The residues M381, K384, E386, and L393 are in close contact with the residues N141, A2 and A4, F21, and R171 of the subdomain IA, respectively, whereas the residue L392 is in close contact with residues G215 and F217 of the subdomain IIA (Figure 3b). In a second trajectory

(APO-2), the initial open model adopted a stable conformation after 70 ns of simulation time, as shown by the time evolution of the C<sup>α</sup> RMSD relative to the initial model, which converged to a constant value (data not shown). Similar events were observed in trajectories APO-1 and APO-2: a break of the initial fused helices A+B, a folding of the SBD-α on the NBD, and the binding of the linker between subdomains IA and IIA. However, in the APO-2 trajectory, helices A and B were found respectively in contact with lobes IA and IB of the NBD, as shown in Figure 3c.

**MD of the Open Model with ADP.** The stable typical structures of the initial open model with ADP (see Methods) relaxed by all-atom MD at 300 K are presented in Figure 3c. In the two trajectories of hHsp70 (ADP-1 and ADP-2, see Methods), we observed a break of the initial fused helix A+B of the initial open model (Figure 1b) and the displacement of the helix bundle toward the NBD (Figure 3c) after 320 ns in ADP-1 and 100 ns in ADP-2, as reflected by the sharp variations of the C<sup>α</sup> RMSD with respect to the initial model at these times (data not shown). The stable positions of the α-helical part of the SBD relative to the NBD are different from the ones obtained for the APO-hHsp70 simulations (Figure 3c). In ADP-1, helices A and B are bound to lobe IA, and in ADP-2, helix A is bound to lobe IA and helix B is bound to lobe IB (Figure 3c).

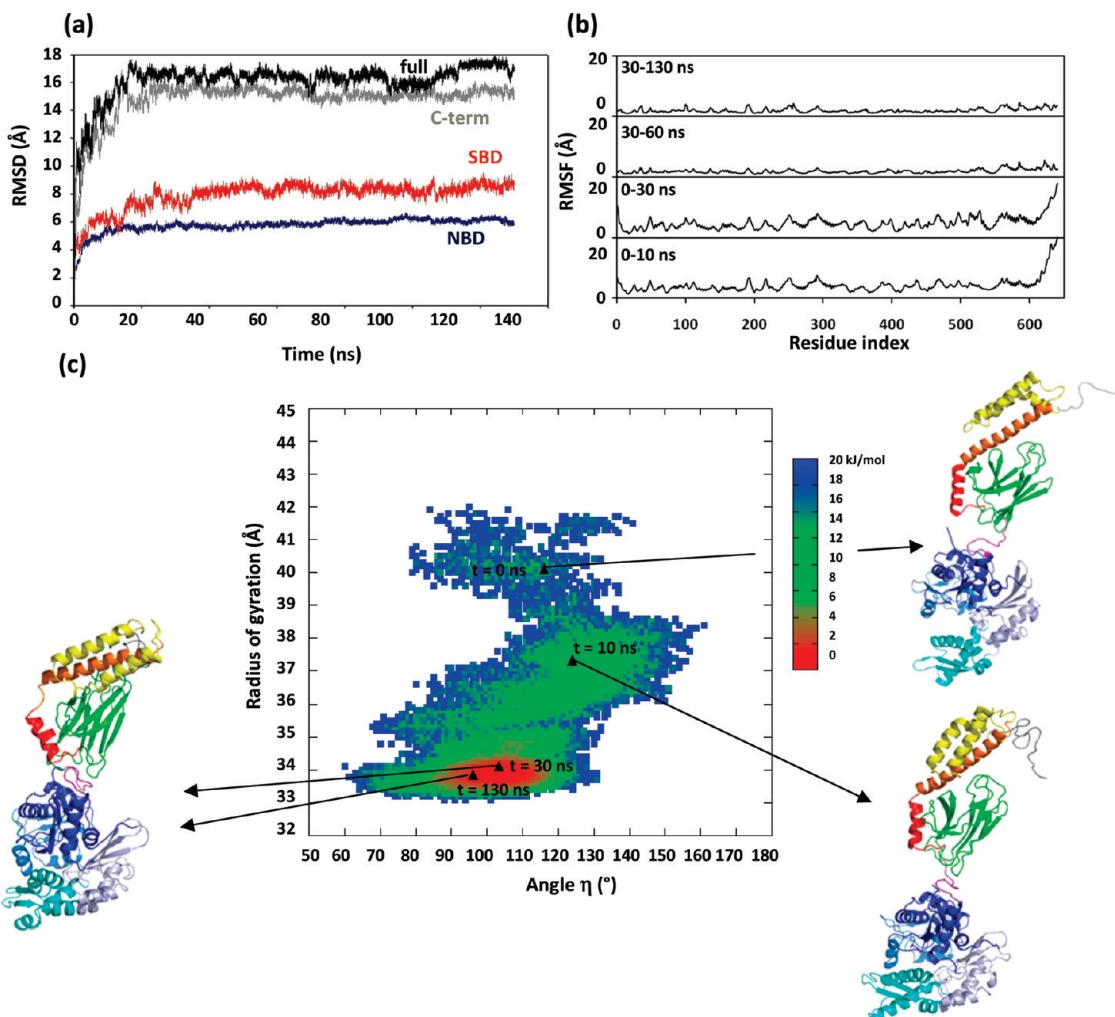
Relatively long all-atom MD simulations are needed to relax the initial model of hHsp70 in an open state (Figure 1b). The model undergoes a major structural relaxation before converging to a stable structure (Figure 3) after several tens of nanoseconds in MD simulations. The structural changes and energy surfaces (not shown) are similar in all trajectories: the fused A+B initial helix breaks into two helices A and B, and the SBD helix bundle moves toward the NBD. The stable average structures of the NBD and of the SBD-β of the open model simulated by MD are similar in the four MD runs: the C<sup>α</sup> RMSD between the four NBD structures varies between 1.9 Å and 4.9 Å, and the C<sup>α</sup> RMSD between the four SBD-β structures varies between 2.5 Å and 3.4 Å. The C<sup>α</sup> RMSD between the SBD-α structures of the four runs varies between 6.4 Å and 10.6 Å because the helices A and B of the SBD-α interact with the NBD, and their precise positions relative to the NBD differ in the different trajectories (Figure 3c). With the NBD of hHsp70 being rather symmetrical, it might be possible for the open state of hHsp70 to exist in different stable conformations involving different positions of the SBD-α relative to the NBD. The structure of the open state of hHsp70, which should correspond to ATP-hHsp70, has not been solved at the atomic level so far.

The hydrolysis of ATP is known to promote substrate binding and facilitate the closure of SBD.<sup>16,17</sup> In the homologous truncated structure solved experimentally,<sup>32</sup> ADP-Hsp70 was found in a closed state (as in the initial model shown in Figure 1c) with the linker exposed to solvent.<sup>32</sup> However, the sole binding of ADP to the initial open model does not provoke the closure of the SBD in our MD simulations (Figure 3c). This might be due to the existence of an activation barrier between the open and closed states that the system cannot cross spontaneously within the

limited time-scale sampled (400 ns) by our MD simulations. This activation barrier may correspond to the extrusion of the linker from its strongly bound position on a hydrophobic cleft between subdomains IA and IIA<sup>22</sup> of the NBD, as found in the initial open model (Figure 1b) and in the stable relaxed APO and ADP structures (Figure 3c). *In vivo*, the transition of hHsp70 between open and closed states involved the hydrolysis of ATP<sup>16,17</sup> and is assisted by cochaperones such as Hsp40 and Hsp110<sup>18–20</sup> not considered in the present study. Simulations of the conformational dynamics of hHsp70 on a longer time-scale ( $\gg\mu$ s) and with cochaperones are beyond the reach of all-atom simulations and will require the application of anharmonic coarse-grained models like UNRES.<sup>58</sup>

**3.2. Structure of Hsp70 Relaxed by MD in a Closed State.** *Initial Model Built by Homology.* The initial model of hHsp70 in a closed state (named the closed model here) was built by homology<sup>48–50</sup> using the recent complete (unfortunately still truncated by 30 residues) structure of DnaK in solution (PDB ID: 2KHO;<sup>32</sup> see Methods). The experimental structure of the bacteria homologue DnaK is locked in the ADP state with a short peptide (NRLLL TG) bound within the SBD-β (PDB ID: 2KHO).<sup>32</sup> The coordinates of the nucleotide and of the peptide substrate were not resolved in the data deposited in the PDB. Compared with the initial open model, the SBD-α of this second model has a kink between the two helices A and B (Figure 1c), as in all known Hsp70 truncated structures.

**MD of the Closed Model without a Nucleotide.** We considered one trajectory of the initial closed model (Figure 1c) without a nucleotide and without a substrate (APO; see Methods). The APO closed model adopted a stable conformation after about 30 ns, as shown by the time evolution of the C<sup>α</sup> RMSD of the full structure with respect to the initial model (Figure 4a) and by the time evolution of the C<sup>α</sup> RMSF of the structure along the sequence (Figure 4b). In Figure 4a, the C<sup>α</sup> RMSD of the full structure and of the SBD converges to a constant value after 30 ns, whereas the C<sup>α</sup> RMSD of the NBD converges after only 10 ns. In Figure 4b, the C<sup>α</sup> RMSF computed between 30 and 60 ns and that between 30 and 130 ns were very similar. Analysis of the C<sup>α</sup> RMSF along the primary sequence in the first 30 ns (Figure 4b) showed that the largest structural fluctuations of the initial model were mainly located in the C-terminal part of the model (residue index  $\geq 616$  in Figure 4b). The C-terminal part of hHsp70 (residues 616–641) was absent in the experimental DnaK template (2KHO) and added in the homology modeling procedure. The largest fluctuations observed in Figure 4b in the first 30 ns were due to the folding of this initial unstructured C-terminal part of the protein. The folded C-terminal part moved between the SBD-β and helix B of the SBD-α, inducing the unfolding of helix E of the initial model (Figure 5a). In addition, the kink between helix A and helix B in SBD-α is more elongated in the relaxed structure than in the initial model, and its SBD-β is rotated (Figure 5a). No other major structural changes were observed after 30 ns up to the end of the MD trajectory (130 ns). The SBD is always closed (Figure 5a) and adopted similar conformations to that



**Figure 4.** Time evolution of structural parameters of the closed model as a function of time. (a) The C<sup>α</sup> RMSD (for the full structure, the NBD, the SBD, and the C-terminal part) computed with respect to the initial closed model built by homology modeling. (b) The C<sup>α</sup> RMSF ( $\sqrt{\langle \Delta R_f^2 \rangle}$ , see eq 4 in Methods) computed for the closed model in different time windows. (c) Energy surface (in kJ/mol) for the relaxation of the closed model studied by MD. The energy surface is defined by  $E = -kT \ln[P(R_g, \eta)/\text{Max}\{P(R_g, \eta)\}]$  where  $P(R_g, \eta)$  is the probability density of the radii of gyration  $R_g$  and of the angle  $\eta$  defined in section 3.1 and computed from MD.  $E$  is zero at the maximum value of  $P(R_g, \eta)$  (red basin).

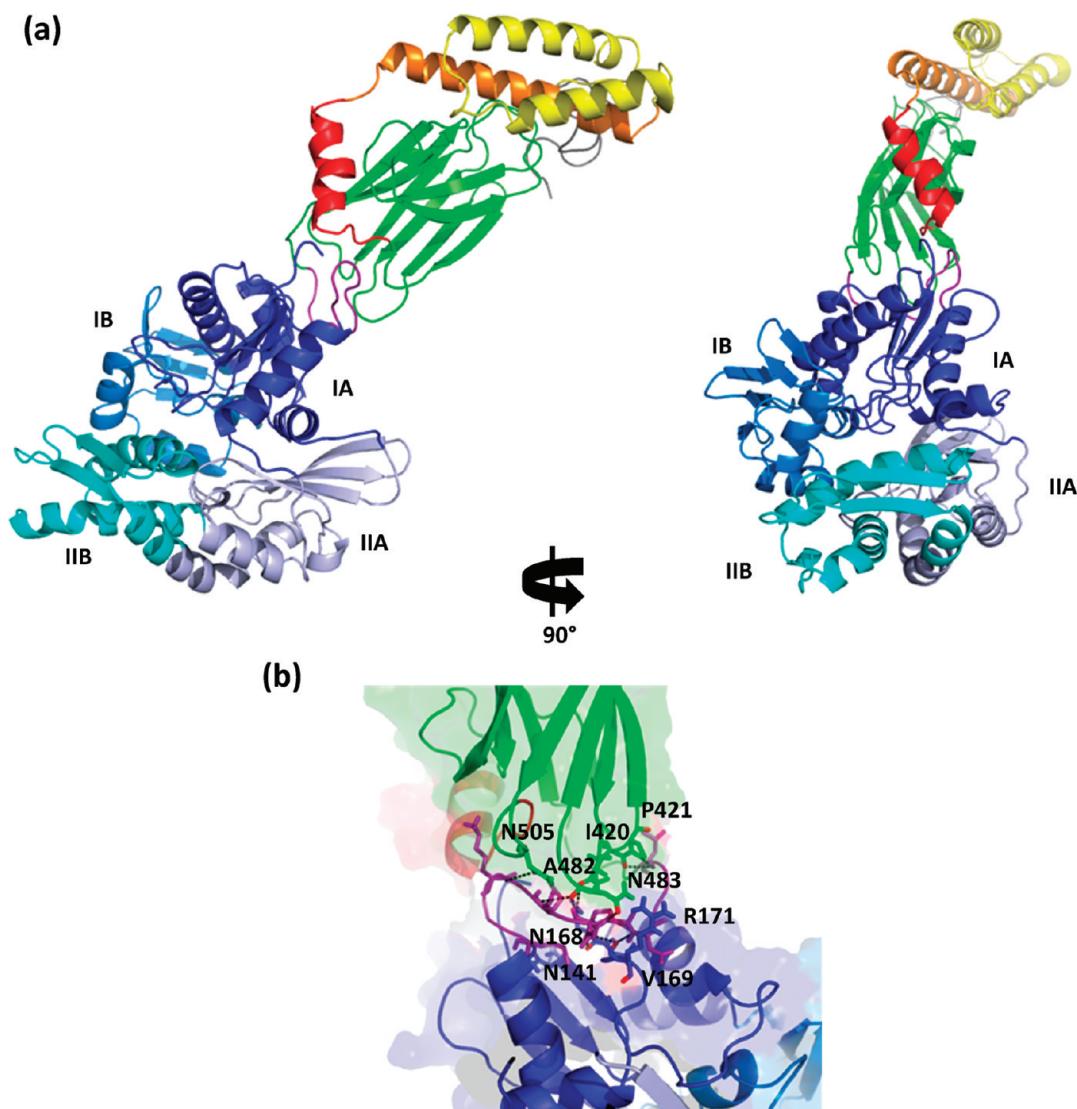
observed in ADP-DnaK.<sup>29,32</sup> On the other hand, preliminary results (unpublished data) for two all-atom MD runs (150 ns) of DnaK in the closed state showed that the bacterial structure adopted stable conformations between 20 and 50 ns which are very similar to that of human Hsp70 in the closed state (Figure 5a). Because the model of the closed state converged quickly to a stable conformation and because the simulations of the closed model of hHsp70 are computationally demanding (the simulation cell contains about 645 000 atoms) and our computational resources are limited, we did not perform additional MD runs.

The sequence of major events of the relaxation of the closed model simulated by MD is shown on the energy surface  $E(R_g, \eta)$  (Figure 4c) where  $R_g$  and  $\eta$  are the radius of gyration of the protein and the angle measuring the curvature of the “A+B” helix as defined in section 3.1. In the first 10 ns, folding of the C-terminal part of the protein was observed. After 30 ns, the SBD and NBD have moved toward each other with the SBD bound to the NBD in a stable conformation, which does not change significantly until the end of

the simulation (130 ns; the RMSD between the structure at 30 ns and the one at 130 ns is 3.4 Å).

The interdomain linker of hHsp70 relaxed by MD in the closed state is not bound in the same hydrophobic cleft between subdomains IA and IIA (Figure 5b), as was the case in the open state of hHsp70 (Figure 3b). However, in the closed state of hHsp70, the linker makes close contacts with the subdomain IA of the NBD; namely, the residues M381, V388, Q389, and L390 of the linker are respectively in close interactions with the residues N141, R171, V169, and N168 of domain IA. In addition, the linker contacts SBD-β: the residues E386, V388, Q389, and L393 of the linker are in close contact with the residues N505, N483, A482 and I420, and P421 of the SBD-β, respectively.

**3.3. Comparison with Experimental Structural Information.** Several structures of two-domain Hsp70 NBD-SBD<sup>31,32</sup> constructs have been published in the past few years. However, all of these structures were truncated, so a comparison of our full-length models with the experimental structures solved by XRD (X-ray diffraction) or NMR



**Figure 5.** Typical structure of the closed model after relaxation by MD simulations. The color code is similar to the one in Figure 1a. (a) Ribbon diagram of the average structure of the closed model in two views rotated by 90°. (b) Close-up view of the contacts between the linker and the other domains. This figure was prepared with PyMOL [http://www.pymol.org].

**Table 1.** Experimental Structural Information for Hsp70 Chaperone and Their Homologues

protein	Uniprot ID	sequence length (aa)	domains	sequence of the constructs (aa)	PDB ID	ligand	exptl. techniques
hHsp70	<b>P08107</b>	641	NBD	3–382	1HJO <sup>26</sup>	ADP	XRD
bHsc70	<b>P19120</b>	650	NBD+SBD $\beta$	1–554	1YUW <sup>31</sup>	ATP	XRD
			NBD	1–386	1NGA <sup>24</sup>	ADP	XRD
			NBD	1–386	1NGG <sup>24</sup>	ATP	XRD
DnaK	<b>P0A6Y8</b>	638	NBD+SBD $\beta+\alpha$	4–603	2KHO <sup>32</sup>	ADP	NMR
Sse1 (Hsp110)	<b>P32589</b>	693	NBD+SBD $\beta+\alpha$	1–649	3C7NA <sup>18</sup>	ATP	XRD

(nuclear magnetic resonance) is not fully appropriate (Table 1). The only experimental data available for a complete bovine Hsc70<sup>36</sup> (650 amino acids) and *E. coli* DnaK<sup>37</sup> (638 amino acids) Hsp70 chaperones are the low-resolution SAXS data.<sup>36,37</sup> Two quantities are typically extracted from SAXS: the radius of gyration  $R_g$  (Table 2) and the probability distribution  $P(r)$  of the distances  $r$  between the heavy atoms (Figure 6).

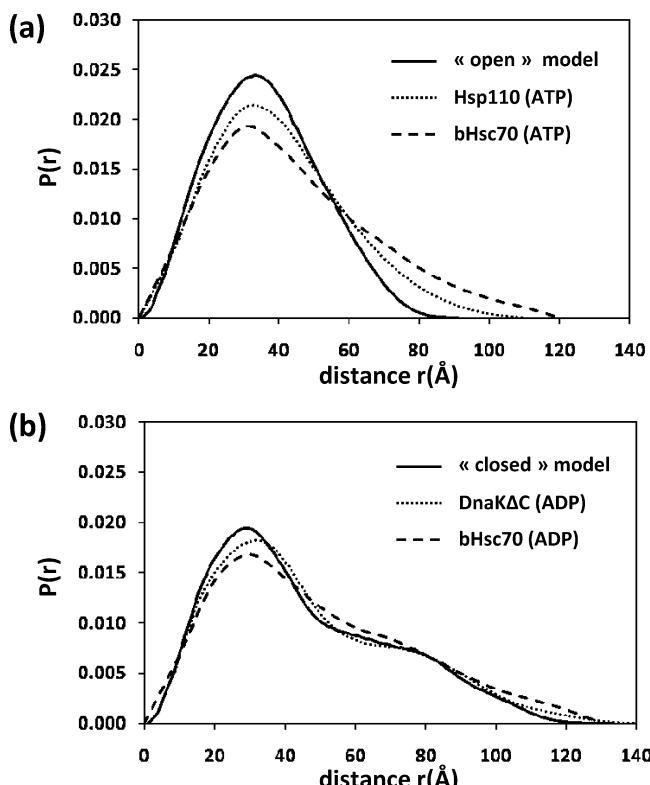
**Radius of Gyration  $R_g$ .** In refs 36 and 37, the radii of gyration  $R_g$  was evaluated from the SAXS data by using both the Guinier approximation [ $\ln(I(Q)/I(0)) = -\exp(-R_g^2 Q^2/3)$ , where  $I(Q)$  is the intensity collected at the (small)

scattering wave-vector  $Q$ ] and by computing numerically the second moment of the probability distribution  $P(r)$  [ $R_g^2 \equiv \int dr r^2 P(r)$ ]. The comparison of the two methods provides an estimation of the actual error bar on  $R_g$ . The difference between the values of the  $R_g$  evaluated by the two methods from the same set of experimental data is about 3.6 Å for ADP-bHSc70 (closed) and 1.9 Å for ATP-bsHc70 (open) (Table 2). On the other hand, for ADP-bHsc70, two identical experiments gave different answers:  $R_g = 34.9$  Å (Guinier) and  $R_g = 37.5$  Å (Guinier), which points to a possible dependence on the protein sample preparation. The largest dispersion of the values of the gyration radius is observed

**Table 2.** Experimental Radii of Gyration  $R_g$  (Å)

protein	ADP			ATP		
	Guinier	$P(r)$	XRD	Guinier	$P(r)$	XRD
bHsc70 <sup>36</sup>	34.9 ± 0.2	38.5 ± 0.5		32.8 ± 0.3	34.7 ± 0.8	
	37.5 ± 1.0	41.7 ± 0.5		33.4 ± 0.2	34.5 ± 0.6	
DnaK <sup>32,37</sup>	45.4 ± 1.5		37.6	39.8 ± 1.0	40.5 ± 1.0	
Sse1 (Hsp110) <sup>18</sup>						31.4
bHsc70 NBD (1–386) <sup>24,36</sup>	22.8 ± 0.2	22.6 ± 0.2	21.4	22.6 ± 0.1	22.2 ± 0.2	21.2
	23.1 ± 0.2	22.9 ± 0.2				
NBD in DnaKΔC <sup>32</sup>			21.1			
NBD in Sse1 <sup>18</sup>						20.5
hHsp70 NBD <sup>26</sup>			21.3			
SBD in DnaK <sup>32</sup>			20.6			
SBD in Sse1 <sup>18</sup>						38.9

for the full structures. For instance, the gyration radius of the full-length DnaK evaluated from the Guinier approximation from SAXS data is 45.4 Å, whereas the gyration radius of the template of the closed model (DnaKΔC) is 37.6 Å from the XRD data (the difference cannot be accounted for by the 30 missing residues in DnaKΔC, since the difference



**Figure 6.** Comparison of the distance distribution functions  $P(r)$  computed for the different models of hHsp70 relaxed by MD with those derived from SAXS data for bHsc70 and those computed for the homologue proteins Hsp110 and DnaK. The function  $P(r)$  of bHsc70 in the ATP and ADP states was digitized from Figure 4 of ref 36. The function  $P(r)$  for the homologue proteins Hsp110 (PDB ID: 3C7N A) in the ATP state and DnaKΔC in the ADP state (PDB ID: 2KHO) were computed from the structures deposited in PDB. (a) Distance distribution function  $P(r)$  of the open relaxed model (apo; solid line), Hsp110 (locked in the ATP state; dotted line), and bHsc70 (locked in the ATP state; dashed line). (b) Distance distribution function  $P(r)$  of the closed relaxed model (apo; solid line), DnaKΔC (locked in the ADP state; dotted line), and the bHsc70 (locked in the ADP state; dashed line).

**Table 3.** Computed Radii of Gyration  $R_g$  (Å)

	MD simulations	hHsp70	NBD in hHsp70	SBD in hHsp70
open model	APO-1	26.4	21.2	28.9
	APO-2	27.5	20.4	26.7
	ADP-1	27.4	20.7	30.3
	ADP-2	27.0	20.5	29.9
closed model	APO	34.1	20.2	19.4

between the radius of gyration of the full length structure of hHsp70 simulated by MD in a closed state and the same structure of hHsp70 with the last 30 residues removed is only 1.1 Å). Taking all of this into consideration, we estimate an error bar on the radius of gyration for the full structure of Hsp70 of at least 3–4 Å. Contrary to our findings for the full structure, the radii of gyration for the isolated NBD of Hsp70 measured by SAXS in solution or by XRD in crystals differ by 1–2 Å (Table 2).

In all of the SAXS experimental data,<sup>36,37</sup> ATP-Hsp70 is found to be systematically more compact than the ADP-Hsp70 state (Table 2). Other experimental data, such as the measurement of tryptophan fluorescence<sup>17,59,60</sup> and proteolytic susceptibility,<sup>17,61</sup> support the hypothesis of a change in protein shape between the ADP and ATP states of Hsp70. The difference in size is due to a change of the radii of gyration of the SBD between the two states. Indeed, the  $R_g$ 's of the isolated NBD of bHsc70 with ATP and ADP are very similar in both SAXS and XRD data (Table 2). The sizes of the NBD isolated and within the crystallized structure of DnaKΔC and that of the full length Hsp110 are also similar. On the other hand, the opening of the SBD increases its radius of gyration by about 20 Å, as we roughly estimated by comparing the size of the SBD within ADP-DnaKΔC and within ATP-Hsp110 (Table 2). The magnitude of the change is probably overestimated because in Hsp110 the “A+B” helix is not broken (Figure 1b).

In agreement with the SAXS data,<sup>36,37</sup> the average stable structures of hHsp70 relaxed by MD in an open state (Figure 3) are more compact than the structure in a closed state (Figure 5). The values of the  $R_g$  computed from MD (Table 3) for the open model varied between 26.5 and 27.5 Å between the four runs (for comparison, the closest value found in SAXS for ATP-bHsc70 is 32.8 Å, Table 2), and the  $R_g$  of the closed model was 34.1 Å (for comparison, the closest value found in SAXS for ADP-bHsc70 is 34.9 Å). The radius of gyration of the NBD domain within the full length hHsp70 model did not vary between the open and

closed states, in agreement with the values of the radii of gyration of ATP-bHsc70 and ADP-bHsc70 (Table 2), and its value is close to the experimental value. The size of the SBD within the closed model of hHsp70 is similar to that of DnaKΔC (Table 3). The opening of the SBD increases the radius of gyration of hHsp70 by about 9–10 Å (Table 3).

**Distance Distribution Function  $P(r)$ .** The function  $P(r)$  of ATP-bHsc70<sup>36</sup> measured by SAXS and computed for the structure of the yeast ATP-Hsp110<sup>18</sup> (PDB ID: 3C7N A), both in an open state, are compared to the  $P(r)$  computed from the open model of hHsp70 relaxed by MD in Figure 6a [ $P(r)$  from run APO-1 is shown; similar results were obtained from the four MD runs of the open model (data not shown)]. The  $P(r)$  functions extracted from SAXS data and from the experimental structure of Hsp110 have one peak at about 35 Å (Figure 6a) which might reflect the binding of NBD on SBD. Indeed, the  $P(r)$  computed for the open model of hHsp70 has a peak at the same position, and the overall shape of  $P(r)$  agrees quite well with the one measured for bHsc70 by SAXS. The ATP-Hsp structures (Figure 6a) are substantially contracted structures by comparison with ADP-Hsp structures (Figure 6b). Indeed, the function  $P(r)$  measured by SAXS for ADP-bHsc70<sup>37</sup> and the function  $P(r)$  calculated from the experimental structure of ADP-DnaKΔC<sup>32</sup> (PDB ID: 2KHO; truncated by 30 residues in the C-terminal part) have two peaks: one at 30 Å and a distinct shoulder at 80 Å (Figure 6b), suggesting a structure with a bilobal shape. The function  $P(r)$  for the closed model of hHsp70 has also a second peak around 80 Å as in the ADP molecular chaperones.

The agreement between theory and experiment is satisfactory taking into account that we compared simulations of human Hsp70 with experimental data of homologue proteins (bovine, bacteria, and yeast). In addition, one must take into account the necessary inaccuracy of any MD force field, the finite length of the MD simulations, and the difference between the idealized solvent simulated (pure water and ions) and the experimental one. Indeed, the experimental conditions can influence the size ( $R_g$ ) and shape [ $P(r)$ ] of the protein due to the fact that the solution contains both ADP and ATP in addition to buffers such as Tris·HCl,<sup>36,37</sup> whereas in our simulation, there are only water molecules and counterions. In addition, as stated by the authors in ref 37, the steady-state solution of ATP-bHsc70 is in fact a mixture of the protein in the open and closed conformations.

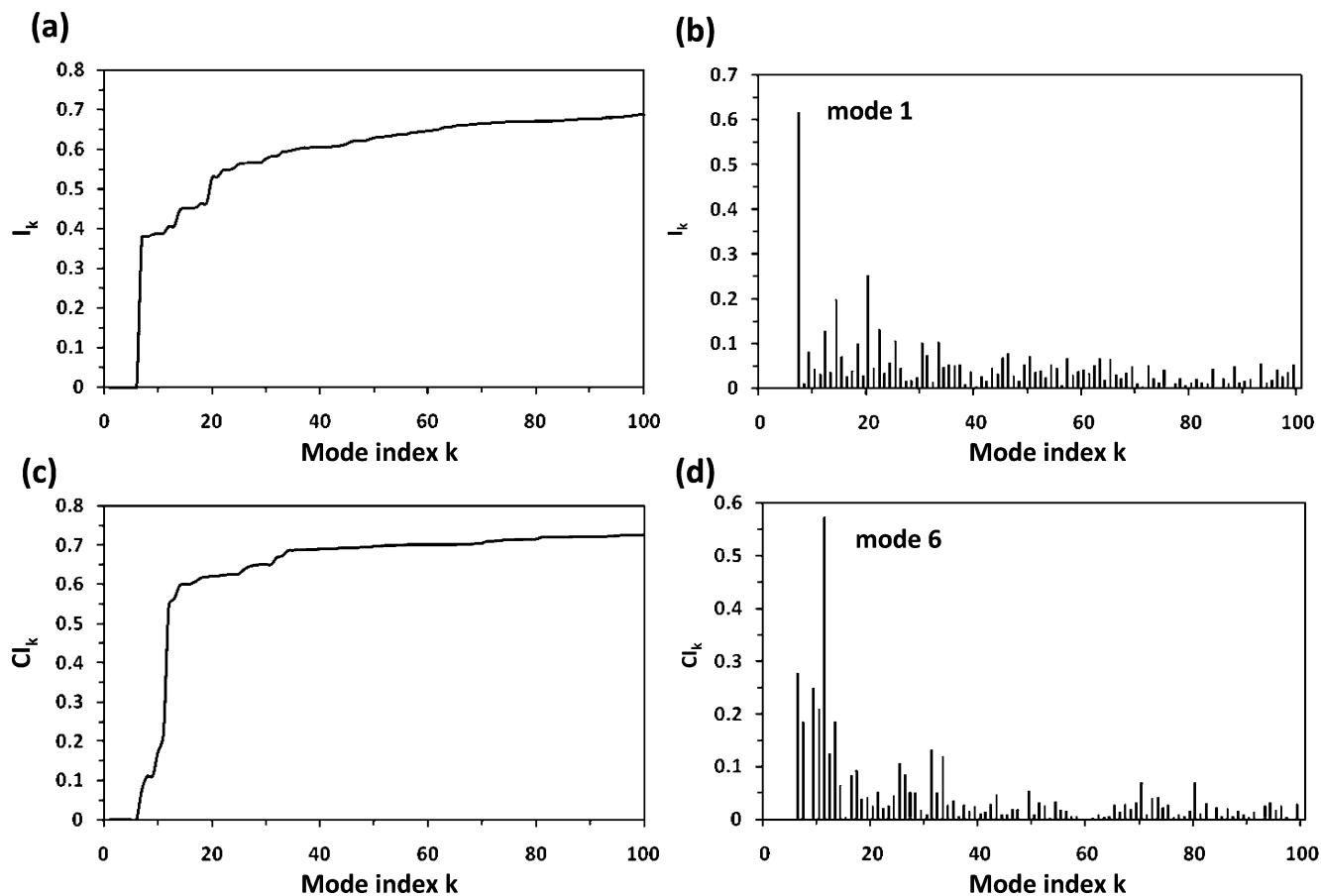
Comparison of the relaxed open (Figure 3) and the closed models (Figure 5) with all experimental data<sup>36,37,59–62</sup> confirms that the former is a good representative model of ATP-Hsp70, whereas the latter is a good representative model of ADP-Hsp70.

**3.4. Conformational Changes between the Open and the Closed Conformations of Hsp70.** The relaxed models of Hsp70 represent initial and final structures of a conformational transition of the functional cycle (see next section) of this chaperone. The conformational change is induced *in vivo* by several factors: nucleotide exchange (ATP↔ADP),<sup>16,17</sup> hydrolysis, and all steps being assisted by cochaperones.<sup>18–20</sup>

In an attempt to identify the functionally important motions for the transition between the open and closed state, we computed the collective modes of hHsp70 by applying the anisotropic network model (ANM)<sup>41–46</sup> to the open model and to the closed model obtained after relaxation by MD (see Methods). As shown for a few proteins, the collective modes of the structures in the initial and final states of a conformational change contain information on the dynamic of the transition.<sup>56,63–68</sup> Because the collective modes have a low restoring force constant and correspond to displacements of large structural subunits of a protein, a simplified elastic model (ANM) is sufficient to reproduce these modes.<sup>41–46</sup>

**Transition from the Open State to the Closed State.** First, we computed the collective modes of the open model and their relevance (involvement coefficient, see Methods) for the transition from the open state to the closed state of hHsp70. The theory is discussed in detail in the Methods section. In brief, a linear pathway interpolating between the two conformations (open and closed) was built. For each collective mode  $k$ , the projection of the atomic displacements within the mode  $k$  on the interpolated pathway defined the involvement coefficient  $I_k$  of the mode (the maximum value is 1, corresponding to a perfect match between the displacements of the atoms within the mode and the interpolating pathway). The sum of the square involvement coefficient of each mode up to an index  $k$  is the cumulative involvement coefficient  $CI_k$  (see Methods). The cumulative involvement coefficient indicates that the first 10 and the first 100 slow modes of a total of the 1917 modes of nonzero frequency of hHsp70 account for 45% ( $CI_{10} = 0.45$ ) and 69% ( $CI_{100} = 0.69$ ) of the displacement from the open to the closed state, respectively (Figure 7a).

The mode contributing the most to the transition between the open state and the closed state is the mode having the lowest nonzero “frequency”  $\lambda_1$  (mode 1), which has an involvement coefficient of 0.62 for the transition from the open state to the closed one (Figure 7b). In this mode, the NBD essentially moves as a rigid unit, whereas the SBD is the most mobile part. The amplitudes of the  $C^\alpha$  B factors (see eq 5 in the Methods) for this mode (Figure 8a) clearly show that only very large fluctuations are present in the SBD-β and in the SBD-α, which correspond to the substrate pocket and the lid, respectively. In SBD-β,  $C^\alpha$  B factors show four peaks for residues 405, 433, 467, and 496 which correspond to important fluctuations in β strands. In SBD-α, the helix A and the C-terminal part are completely rigid (residues 508–525 for helix A and residues 616 to 641 for the C-terminal part in Figure 8a), and helix A forms a hinge region. The global motion described by the lowest frequency mode shows that the helix A of SBD-α serves as a hinge region around which SBD-β and the rest of SBD-α (helix B + C + D) move toward (Figure 9). SBD-β and SBD-α move in opposite directions from each other (Figure 9), and  $R_g$  is nearly constant (Figure 10a). For all of the structures of the open model relaxed by all-atom MD (runs APO-1, APO-2, ADP-1, and ADP-2), the transition from the open state to the closed state was typically described by the same collective mode (Figure 9): we observed the rotation of the



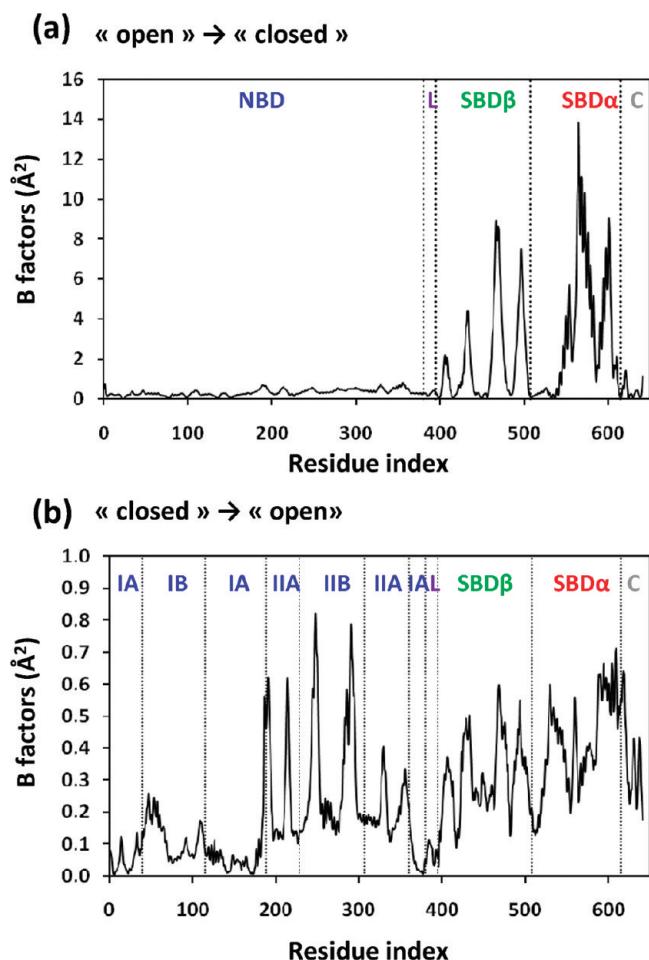
**Figure 7.** Individual and cumulative involvement coefficients (eqs 9 and 10 in Methods). The individual (a) and cumulative involvement coefficients (c) of the first 100 modes of nonzero frequency for the transition from the open state to the closed state of human Hsp70. The modes are computed for the average structure of the open model calculated between 100 and 400 ns of the MD run APO-1. The individual (b) and cumulative involvement coefficients (d) of the first 100 modes for the transition from the closed state to the open state of human Hsp70. The modes are computed for the average structure of the closed model calculated between 30 and 130 ns of the MD run.

SBD- $\alpha$  in the direction of the substrate binding pocket, which simulates the closing of the pocket by the lid.

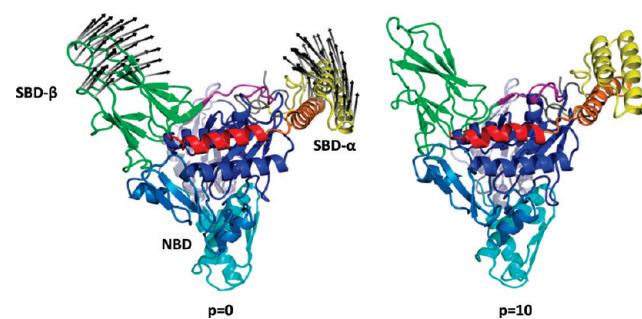
*Transition from the Closed State to the Open State.* Second, we computed the modes for the relaxed closed model in order to study the transition from the closed to the open state. The first 10 and the first 100 collective modes of a total 1917 nonzero frequency modes of hHsp70 account for 60% ( $CI_{10} = 0.6$ ) and 73% ( $CI_{100} = 0.73$ ) of the displacement from the closed to the open state, respectively (Figure 7c). In the closed state, the mode contributing the most to the transition between the closed state and the open state is the mode having the lowest “frequency”  $\lambda_6$  (mode 6), which has an involvement coefficient of 0.57 (Figure 7d). This motion tends to decrease ( $p > 0$ ) or to increase ( $p < 0$ ) the radius of gyration of the protein (Figure 10a). Indeed, this motion corresponds to a compression of the two domains ( $p > 0$ , Figure 11a), which restricts the mobility of the linker, or to an elongation of the two domains [ $p < 0$ , not shown]. This mode does not correspond to a direct opening of the lid although there are important fluctuations in the SBD as shown by the  $B$  factors computed for this mode (Figure 8b).

The analysis of the  $B$  factors shows clearly that mode 6 involves displacements of residues both in the NBD and in the SBD (Figure 8b). In particular, residues 190, 214,

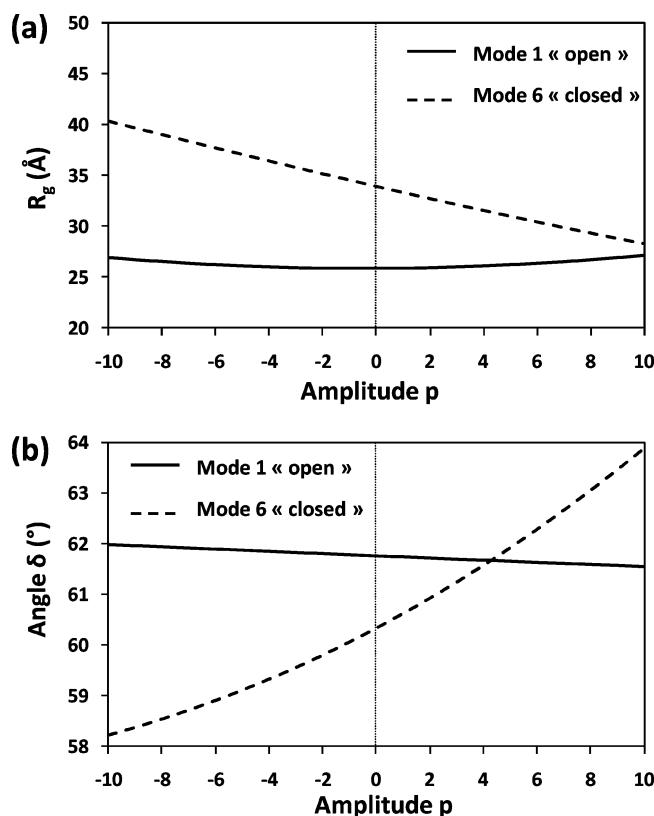
247–249, and 289–291 in the NBD are very mobile. These residues belong to subdomain II: the most mobile residues (residues 247–248, 289–291) are in subdomain IIB, and the others (residues 190, 214) in subdomain IIA. The ribbon diagrams displayed in Figure 11b show the displacement of subdomain IIB in the closed to open transition. The large  $B$  factors of residues 190, 214, 247–249, and 289–291 are due to the motion of subdomain IIB of the NBD, which rotates and opens slightly the nucleotide pocket in NBD (Figure 11b). We quantified the opening of the NBD by computing the angle  $\delta$  formed by the centers of mass of subdomains IB, IA/IIA, and IIB.<sup>69</sup> The angle  $\delta$  increased (for  $p > 0$ ) for mode 6, confirming the opening of NBD (Figure 10b). Contrarily, in the transition from the open state to the closed state of hHsp70, the angle  $\delta$  did not vary for mode 1 of the open structure of hHsp70, as shown in Figure 10b. On the other hand, the average value of angle  $\delta$  is larger when a nucleotide is bound to the NBD: the average value of  $\delta$  of APO-hHsp70 was a few degrees larger than the average value of  $\delta$  of ADP-hHsp70 in our MD simulations. In recent MD simulations of isolated NBD of hHsp70, angle  $\delta$  of a nucleotide-free NBD was also found to be 10° larger than that of an ATP-NBD.<sup>69</sup>



**Figure 8.**  $B$  factors ( $\text{\AA}^2$ ) computed at 300 K [eq 5 in the Methods]. (a) Computed  $B$  factors for the collective mode 1 of the open model of the hHsp70 chaperone having the largest involvement coefficient in the transition from the open state to the closed state (Figure 7b). (b) Computed  $B$  factors for the sixth slowest mode, 6, of the closed model of hHsp70 having the largest involvement coefficient in the transition from the closed state to the open state (Figure 7d). The dotted lines show the different subdomains of hHsp70 in agreement with Figure 1.



**Figure 9.** Graphical representation of the collective mode  $\lambda_1$  of the open model of hHsp70 chaperone. The color code is the same as in Figure 1a. Eigenvectors are represented by black arrows. Vector norms have been uniformly scaled by an arbitrary factor, and only vectors with the higher norms are represented for clarity. The quantity “ $p$ ” represents the amplitude of the mode in arbitrary units. This figure was prepared with PyMOL [<http://www.pymol.org>].

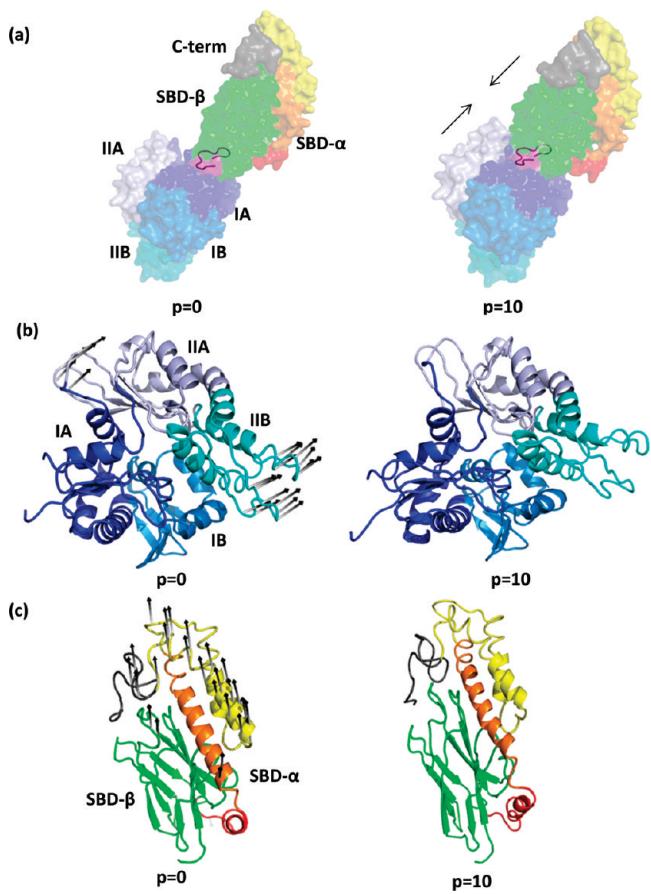


**Figure 10.** (a) Evolution of the radius of gyration ( $R_g$ ) of the structure in the function of the amplitude of mode 1 (open → closed; solid line) and of mode 6 (closed → open; dashed line). (b) Evolution of the opening angle  $\delta$  of the NBD (see text) as a function of the amplitude of modes 1 and 6. The representation is the same as in panel a.

The large  $B$  factors of residues 530 and 608–609 in mode 6 (Figure 8b) are due to motion that involves the displacement of the helix bundle (C + D) and the C-terminal part moving in the opposite direction of SBD- $\beta$  (Figure 11c).

Three other modes contribute significantly to the transition from the closed to the open state (Figure 7d). The first ( $\lambda_1$ ), the fourth ( $\lambda_4$ ), and the fifth ( $\lambda_5$ ) lowest frequency modes have involvement coefficients of 0.28, 0.25, and 0.21, respectively (Figure 7d). In the first mode ( $\lambda_1$ ), the NBD and the SBD move toward each other with a “scissor” movement, in which the linker is a rigid unit (data not shown). The most mobile part in the NBD is subdomain IIA, whereas in the SBD, the SBD- $\alpha$  and the C-terminal part are the most mobile. The fourth mode ( $\lambda_4$ ) corresponds to an opening of the SBD of the protein (data not shown). The most mobile part in the SBD is the SBD- $\alpha$  and the C-terminal part. Finally, in the fifth mode ( $\lambda_5$ ), the NBD and the SBD move toward each other with an important motion in the two domains (data not shown): in the NBD, lobe II is the most mobile part, and the motion in the SBD corresponds to an opening of the lid with huge displacements in the SBD- $\alpha$ . These three modes have in common a motion within the SBD and a coupling between the motion of the SBD and the motion of the NBD, principally via the rotation of lobe II of the NBD.

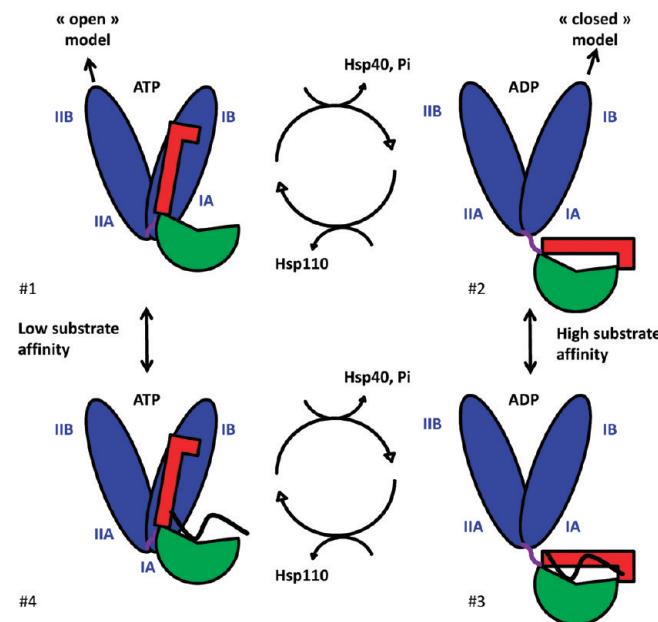
The study of the transition from the closed state to each of the different other structures of the open model (runs APO-



**Figure 11.** Graphical representation of mode 6 of the closed model of the hHsp70 chaperone. The color code is similar to the one in Figure 1a. (a) Surface diagram of the hHsp70 chaperone. The linker is also represented with a ribbon. (b) Ribbon diagram of the NBD. Eigenvectors are represented by black arrows. Vector norms have been uniformly scaled by an arbitrary factor, and only vectors with higher norms are represented for clarity. (c) Ribbon diagram of the SBD. The vectors representation is the same as in panel b. The quantity “p” represents the amplitude of the mode in arbitrary units. This figure was prepared with PyMOL [<http://www.pymol.org>]

1, APO-2, ADP-1, and ADP-2) gave similar results (data not shown). For each run, the sixth lowest frequency mode contributes the most to the transition with the same motion shown in Figure 11. One emphasizes that the collective modes involved in the transition from the open state to the closed state and from the closed state to the open state are very different (Figures 8–11) as well as the dependence of the cumulative involvement coefficient with the mode index (Figure 7).

**3.5. Hsp70 Chaperone Cycle and Interdomain Communication.** The two nucleotides have similar binding energy to Hsp70; the affinity of ATP is slightly higher [the (apparent) Gibbs free energies at 25 °C of the binding of ATP and ADP to a mutant of Dnak lacking hydrolysis activity are  $\Delta G = -39$  kJ/mol and  $\Delta G = -37.7$  kJ/mol, respectively].<sup>70</sup> The binding of ATP to a nucleotide free Hsp70 induces a conformational change.<sup>17,37,60</sup> The hydrolysis of the bound ATP converts slowly ATP-Hsp70 into ADP-Hsp70.<sup>17,70</sup> An exchange reaction between the products



**Figure 12.** Schematic representation of the Hsp70 cycle. The color code is as follows: NBD = blue, SBD-β = green, SBD-α = red, the so-called linker = purple, and the substrate peptide in black. The two models of human Hsp70 were built by homology from the experimental data of the proteins (PDB ID: 3C7N, open model; 2KHO, closed model) and relaxed by MD, corresponding to steps 1 and 2 in the cycle, as indicated by arrows. Adapted with permission from ref 22, Figure 6a. Copyright 2007 Cell Press.

of hydrolysis (ADP and inorganic P) and ATP in solution reinitializes the reaction. The rate of hydrolysis and the rate of the nucleotide exchange are increased by several orders of magnitude by cochaperones,<sup>16</sup> and the binding of the polypeptide to ATP-Hsp70 stimulates the ATP hydrolysis.<sup>17</sup>

A detailed mechanistic model of these different steps of the ATPase activity and of the chaperoning cycle (Figure 12) is unknown despite important insights obtained recently from the structures of Hsp70 fragments and complexes of different species. Due to the large number of Hsp70 species, states, and complexes, studied by different experimental techniques, a qualitative scenario of the allosteric mechanism has been proposed.<sup>17,18,21,22,71</sup> However, the absence of structures of Hsp70+ATP leaves unanswered the mechanism by which the protein substrate triggers the conformational change observed by SAXS<sup>36,37</sup> and depicted by the elongation of a linker between the NBD and SBD in moving from steps 1 to 2 or 4 to 3 (Figure 12). Therefore, our two models and simulations of the conformational dynamic of human Hsp70 may provide important insights about transition pathways that experimental data could not reveal so far.

In the present paper, we consider only the chaperoning cycle without a substrate and without cochaperones corresponding to steps 1 and 2 in Figure 12. In ADP-Hsp70 without a substrate (step 2 in Figure 12), represented by hHsp70 in a closed state (Figure 5), the linker is bound to SBD-β. Consequently, the SBD and the NBD of hHsp70 are not highly bound and move rather freely: in our MD simulations, the lowest frequency collective modes of the closed state correspond to different global motions of the

NBD relative to the SBD. The mobility of the NBD relative to the SBD is in agreement with NMR data for ADP-DnaK.<sup>22</sup> When ADP is exchanged by ATP, from step 2 (ADP-Hsp70) to step 1 (ATP-Hsp70), the protein undergoes large conformational changes in both domains,<sup>16,17,22,36,37</sup> and the linker is sequestered from the solvent and binds in the IA/IIA surface cleft as we found for hHsp70 in the open state (Figure 3b). In the open state, we found that the SBD and the NBD are highly bound. The lowest frequency mode of the open state corresponds to oscillation of the substrate binding cleft with little variation of the structure size (Figure 10a).

The collective modes help us to understand the transition from step 2 to step 1 and the interdomain communication. Indeed, we observe a rotation of subdomains IIA and IIB in the NBD (Figure 11b) as well as a rotation of the SBD- $\alpha$  (Figure 11c) in mode  $\lambda_6$ . These two motions suggest that the transition from step 2 to 1 (Figure 12) is assisted by a collective motion which couples a conformational change in the ATPase domain of Hsp70 to the polypeptide binding site.

The rotation of subdomains IIA and IIB in the collective mode  $\lambda_6$  driving the transition from step 2 to step 1 agrees with various experimental data.<sup>22,71</sup> Rotations in the subdomain IIB have been suggested to be the allosteric mechanism in Hsp70 as NMR measurement of truncated structures of bacteria Hsp70<sup>60,71</sup> reveal a rotation of this subdomain intrinsically or after nucleotide exchange. Several nucleotide exchange factors (cochaperones Bag-1 and Hsp110) induce a rotation of IIB upon binding to ADP-Hsp70.<sup>18–20</sup> In particular, subdomain IIB is a key domain of the interaction between Hsp70 and the nucleotide exchange factor (NEF) and cochaperone Hsp110.<sup>18–20</sup> The subdomains IIB, IIA, and IB of the NBD of Hsp110 make contact with subdomains IB, IA, and IIB of Hsp70, respectively. Moreover, the helix “A+B” of the SBD- $\alpha$  of Hsp110 (Figure 1b) contacts subdomain IIB of Hsp70. Therefore, subdomain IIB of Hsp70 is clamped between helix “A+B” of the SBD- $\alpha$  and subdomain IB of Hsp110, and a rotation by 20° allows the exchange of ADP by ATP.<sup>18</sup> The displacement of subdomain IIA in the collective mode from closed to open states (Figure 11b) is also of great interest,<sup>71</sup> since it affects a hydrophobic surface cleft between subdomains IA and IIA, which is likely involved in the binding of the linker of Hsp70 to its NBD.<sup>21,22,72</sup> The present study indicates that the rotation of subdomain IIB is excited thermally even in the absence of a cochaperone. Consequently, the slow modes which described the transition from the ADP (step 2) to the ATP state (step 1) in the chaperone cycle provide a detailed mechanistic description of the allosteric mechanism, in agreement with NMR data for isolated NBD domains of DnaK.<sup>71</sup> In the opposite transition from step 1 to step 2, there is a conformational change in the SBD, which is the closing of the substrate binding pocket by the lid. In the present study, the thermal excitation of mode  $\lambda_1$  and the binding of ADP to the open model of hHsp70 do not generate a force large enough to displace the linker from the buried position adopted during the open state. This suggests that binding of the peptide substrate to the open state and hydrolysis might be necessary to close the SBD. On the other hand, the binding

of the polypeptide substrate could be coupled to the collective mode  $\lambda_1$  which closes the SBD pocket (Figure 9).

## 4. Conclusion

The purpose of the present paper was to obtain stable atomic models of hHsp70 in different states in the absence of cochaperones. Two typical structures corresponding to two steps (1 and 2) of the chaperoning cycle (Figure 12) were reported here. They correspond to human Hsp70, in an open state and a closed state, respectively. The full-length closed and open states models were built by homology modeling<sup>47–50</sup> using as templates the experimental structures of DnaK<sup>32</sup> (closed model; PDB ID: 2KHO) and the homologous protein Hsp110<sup>18</sup> (open model; PDB ID: 3C7N, chain A). The initial models were relaxed by using all-atom MD simulations at 300 K in explicit solvent (see Methods). We obtained stable structures (with and without ADP) by using four different initial conditions for the open state of hHsp70. In the four MD trajectories, the open model adopts a stable conformation in which the lid is open (Figure 3). Different stable binding sites were found for the SBD- $\alpha$  on the NBD, suggesting that hHsp70 in an open state might exist in different conformations. A comparison with experimental data, particularly SAXS<sup>36,37</sup> (Figure 6), suggests that the relaxed open model corresponds to an ATP state of the chaperone cycle. The closed model (2KHO template)<sup>32</sup> was relaxed by all-atom MD simulations at 300 K in explicit solvent (see Methods). The stable structure found by MD has the SBD- $\alpha$  closing the substrate binding pocket located in the SBD- $\beta$ ; this stable structure represents an ADP state of the hHsp70 chaperone.<sup>32</sup>

By assuming that the open state of hHsp70 represents a reasonable model for the ATP-hHsp70 state, we can suggest an explanation of the interdomain communication. The present study of collective modes involved in the transition from step 1 to step 2 of the chaperone cycle (Figure 12) and vice versa reveals one dynamical domain of Hsp70: the SBD. The first lowest frequency mode of the structure in an open state conformation contributing the most to the transition from 1 to 2 of the cycle (Figure 12) corresponds to a closing of the SBD- $\beta$  by the SBD- $\alpha$  (Figure 9). The study of the transition from step 2 to step 1 of the Hsp70 cycle (Figure 12) is dominated by one mode of the structure in a closed state conformation which involved both the NBD and the SBD (Figures 8 and 11). The opening of the substrate binding pocket in this mode causes simultaneously a rotation of subdomains IIA and IIB in the NBD; therefore, it means that this mode provides a means for distant residues in the NBD and SBD to communicate at a low energy cost (weak restoring forces) in the transition from the closed to the open state of human Hsp70. The present calculations demonstrated, for the first time, that the rotation of subdomains IIA and IIB is thus related to the opening of the SBD- $\alpha$  and to an oscillation of the length structure ( $R_g$ ).

**Acknowledgment.** P.S. thanks the Centre National de la Recherche Scientifique (CNRS) for support (délégation CNRS). This research was conducted using the computational resources of the “Centre de Ressources Informatiques” (CRI), Université de Bourgogne. This work was granted to the HPC

ressources of CINES under the allocation 2009-c2009076161 made by GENCI (Grand Equipement National de Calcul Intensif). A.N. thanks the French Ministry of Research and Education for a Ph.D. fellowship. The authors thank Thomas Gaillard for providing a PYMOL script to represent the normal modes.

## References

- (1) Hartl, F. U.; Hayer-Hartl, M. Molecular chaperones in the cytosol, from nascent chain to folded protein. *Science* **2002**, *295*, 1852–1858.
- (2) Young, J. C.; Agashe, V. R.; Siegers, K.; Hartl, F. U. Pathways of chaperone mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 781–791.
- (3) Bukau, B.; Deuerling, E.; Pfund, C.; Craig, E. A. Getting newly synthesized proteins into shape. *Cell* **2000**, *101*, 119–122.
- (4) Saibil, H. R. Chaperone machines in action. *Curr. Opin. Struct. Biol.* **2008**, *18*, 35–42.
- (5) Mayer, M. P.; Bukau, B. Hsp70 chaperones, cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* **2005**, *62*, 670–684.
- (6) Parsell, D. A.; Lindquist, S. The function of heat-shock proteins in stress tolerance, degradation and reactivation of damaged proteins. *Annu. Rev. Genet.* **1993**, *27*, 437–496.
- (7) Qian, S.; McDonough, B.; Boellmann, H. F.; Cyr, D. M.; Patterson, C. CHIP-mediated stress recovery by sequential ubiquitination of substrates and Hsp70. *Nature* **2006**, *440*, 551–555.
- (8) Young, J. C.; Barral, J. M.; Hartl, F. U. More than folding localized functions of cytosolic chaperones. *Trends Biochem. Sci.* **2003**, *28*, 541–547.
- (9) Broadley, S. A.; Hartl, F. U. The role of molecular chaperones in human misfolding diseases. *FEBS Lett.* **2009**, *583*, 2647–2653.
- (10) Warrick, J. M.; Chan, H. Y. E.; Gray-Board, G. L.; Chai, Y.; Paulson, H. L.; Bonini, N. M. Suppression of polyglutamine-mediated neurodegeneration in drosophila by the molecular chaperone Hsp70. *Nat. Genet.* **1999**, *23*, 425–428.
- (11) Muchowski, P. J.; Schaffar, G.; Sittler, A.; Wanker, E. E.; Hayer-Hartl, M.; Hartl, F. U. Hsp70 and Hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7841–7846.
- (12) Jana, N. R.; Tanaka, M.; Wang, G.; Nukina, N. Polyglutamine length-dependent interaction of Hsp40 and Hsp70 family chaperones with truncated N-terminal huntingtin, their role in suppression of aggregation and cellular toxicity. *Hum. Mol. Genet.* **2000**, *9*, 2009–2018.
- (13) Kim, S.; Nollen, E. A.; Kitagawa, K.; Bindokas, V. P.; Morimoto, R. I. Polyglutamine protein aggregates are dynamic. *Nat. Cell Biol.* **2002**, *4*, 826–831.
- (14) Javid, B.; MacAry, P. A.; Lehner, P. J. Structure and function, Heat Shock Proteins and adaptative immunity. *J. Immunol.* **2007**, *179*, 2036–2040.
- (15) Schmitt, E.; Gehrmann, M.; Brunet, M.; Multhoff, G.; Garrido, C. Intracellular and extracellular functions of heat shock proteins, repercussions in cancer therapy. *J. Leukocyte Biol.* **2007**, *81*, 15–27.
- (16) Brehmer, D.; Rudiger, S.; Gassler, C. S.; Klostermeier, D.; Packschies, L.; Reinstein, J.; Mayer, M. P.; Bukau, B. Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. *Nature* **2001**, *8*, 427–432.
- (17) Buchberger, A.; Theyssen, H.; Schröder, H.; McCarty, J. S.; Virgallita, G.; Milkereit, P.; Reinstein, J.; Bukau, B. Nucleotide-induced conformational changes in the ATPase and substrate binding domains of the DnaK chaperone provide evidence for interdomain communication. *J. Biol. Chem.* **1995**, *270*, 16903–16910.
- (18) Schuermann, P. J.; Jiang, J. W.; Cuellar, J.; Llorca, O.; Wang, L. P.; Gimenez, L. E.; Jin, S. P.; Taylor, A. B.; Demeler, B.; Morano, K. A.; Hart, P. J.; Valpuesta, J. M.; Lafer, E. M.; Sousa, R. Structure of the Hsp110:Hsc70 nucleotide exchange machine. *Mol. Cell* **2008**, *31*, 232–243.
- (19) Polier, S.; Dragovic, Z.; Hartl, F. U.; Bracher, A. Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. *Cell* **2008**, *133*, 1068–1079.
- (20) Liu, Q.; Hendrickson, W. A. Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. *Cell* **2007**, *131*, 106–120.
- (21) Vogel, M.; Mayer, M. P.; Bukau, B. Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. *J. Biol. Chem.* **2006**, *281*, 38705–38711.
- (22) Swain, J. F.; Dinler, G.; Sivendran, R.; Montgomery, D. L.; Stotz, M.; Giersch, L. M. Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Mol. Cell* **2007**, *26*, 27–39.
- (23) Flaherty, K. M.; DeLuca-Flaherty, C.; McKay, D. B. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* **1990**, *346*, 623–628.
- (24) Flaherty, K. M.; Wilbanks, S. M.; DeLuca-Flaherty, C.; McKay, D. B. Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. *J. Biol. Chem.* **1994**, *269*, 12899–12907.
- (25) Wisniewska, M.; Karlberg, T.; Lehtio, L.; Johansson, I.; Kotenyova, T.; Moche, M.; Schüler, H. Crystal structures of the ATPase domains of four human Hsp70 isoforms: HSPA1L/Hsp70-hom, HSPA2/Hsp70-2, HSPA6/Hsp70B $\mu$ , and HSPA5/BiP/GRP78. *PLoS ONE* **2010**, *5* (1), e8625.
- (26) Osipiuk, J.; Freeman, C.; Morimoto, R. I.; Joachimiak, A. Structure of a new crystal form of human Hsp70 ATPase domain. *Acta Crystallogr.* **1999**, *D55*, 1105–1107.
- (27) Wang, H.; Kurochkin, V.; Pang, Y.; Hu, W.; Flynn, G. C.; Zuiderweg, E. R. P. NMR solution structure of the 21 kDa chaperone protein DnaK substrate binding domain, a preview of chaperone-protein interaction. *Biochem.* **1998**, *37*, 7929–7940.
- (28) Chou, C. C.; Forouhar, F.; Yeh, Y. H.; Shr, H. L.; Wang, C.; Hsiao, C. D. Crystal structure of the C-terminal 10-kDa subdomain of Hsc70. *J. Biol. Chem.* **2003**, *278*, 30311–30317.
- (29) Zhu, X.; Zhao, X.; Burkholder, W. F.; Gragerov, A.; Ogata, C. M.; Gottesman, M. E.; Hendrickson, W. A. Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* **1996**, *272*, 1606–1614.
- (30) Morshauer, R. C.; Hu, W.; Wang, H.; Pang, Y.; Flynn, G. C.; Zuiderweg, E. R. P. High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70. *J. Mol. Biol.* **1999**, *289*, 1387–1403.
- (31) Jiang, J.; Prasad, K.; Lafer, E. M.; Sousa, R. Structural basis of interdomain communication in the Hsc70 chaperone. *Mol. Cell* **2005**, *20*, 513–524.

- (32) Bertelsen, E. B.; Chang, L.; Gestwicki, J. E.; Zuiderweg, E. R. P. Solution conformation of wild-type *E. coli* Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 8471–8476.
- (33) Schmitt, E.; Maingret, L.; Puig, P. E.; Ghiringhelli, F.; Hammann, A.; Solary, E.; Kroemer, G.; Garrido, C. Hsp70 neutralization exerts potent anti-tumor effects in animal models of colon cancer and melanoma. *Cancer Res.* **2006**, *66*, 4191–4197.
- (34) Didelot, C.; Lanneau, D.; Brunet, M.; Joly, A. L.; De Thonel, A.; Ghiosis, G.; Garrido, C. Anti-cancer therapeutic approaches based on intracellular and extracellular heat shock proteins. *Curr. Med. Chem.* **2007**, *14*, 1–9.
- (35) Leu, J. I. J.; Pimkina, J.; Frank, A.; Murphy, M. E.; George, D. L. A small molecular inhibitor of inducible Heat shock protein 70. *Mol. Cell* **2009**, *36*, 15–27.
- (36) Wilbanks, S. M.; Chen, L.; Tsuruta, H.; Hodgson, K. O.; McKay, D. B. Solution small-angle x-ray scattering study of the molecular chaperone Hsc70 and its subfragment. *Biochem.* **1995**, *34*, 12095–12106.
- (37) Shi, L.; Kataoka, M.; Fink, A. L. Conformational characterization of DnaK and its complexes by small-angle x-ray scattering. *Biochem.* **1996**, *35*, 3297–3308.
- (38) Hinsen, K. Analysis of domain motions by approximate normal mode calculations. *Proteins* **1998**, *33*, 417–429.
- (39) Kitao, A.; Go, N. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.* **2000**, *9*, 164–169.
- (40) Tirion, M. M. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* **1996**, *77*, 1905–1908.
- (41) Atilgan, A. R.; Durell, S. R.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **2001**, *80*, 505–515.
- (42) Yang, L.; Song, G.; Jernigan, R. L. How well can we understand large-scale protein motions using normal modes of elastic network models. *Biophys. J.* **2007**, *93*, 920–929.
- (43) Kim, M. K.; Jernigan, R. L.; Chirikjian, G. S. Efficient generation of feasible pathways for protein conformational transitions. *Biophys. J.* **2002**, *83*, 1620–1630.
- (44) Kim, M. K.; Jernigan, R. L.; Chirikjian, G. S. Rigid-cluster models of conformational transitions in macromolecular machines and assemblies. *Biophys. J.* **2005**, *89*, 43–55.
- (45) Yang, L.; Song, G.; Jernigan, R. L. Protein elastic network models and the ranges of cooperativity. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 12347–12352.
- (46) Navizet, I.; Lavery, R.; Jernigan, R. L. Myosin flexibility: structural domains and collective vibrations. *Proteins* **2004**, *54*, 384–393.
- (47) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST & PSI-BLAST, a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
- (48) Sali, A.; Blundell, T. L. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
- (49) Sali, A.; Potterton, L.; Yuan, F.; Vanvlijmen, H.; Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins* **1995**, *23*, 318–326.
- (50) Sanchez, R.; Sali, A. Comparative protein structure modeling. Introduction and practical examples with MODELLER. *Methods Mol. Biol.* **2000**, *143*, 97–129.
- (51) Shindyalov, I. N.; Bourne, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **1998**, *11*, 739–747.
- (52) Lindahl, E.; Hess, B.; van der Spoel, D. Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **2001**, *7*, 306–317.
- (53) Schuettkopf, A. W.; Van Aalten, D. M. F. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr.* **2004**, *D60*, 1355–1363.
- (54) Ben-Avraham, D. Vibrational normal-mode spectrum of globular proteins. *Phys. Rev. B* **1993**, *47*, 14559–14560.
- (55) Svanidze, A. V.; Sashin, I. L.; Lushnikov, S. G.; Gvasaliya, S. N.; Turoverov, K. K.; Kuznetsova, I. M.; Kojima, S. Inelastic incoherent neutron scattering in some proteins. *Ferroelectrics* **2007**, *348*, 556–562.
- (56) Tama, F.; Sanejouand, Y. H. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* **2001**, *14*, 1–6.
- (57) Baba, A.; Komatsuzaki, T. Construction of effective free energy landscape from single-molecule time series. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19297–19302.
- (58) Maisuradze, G. G.; Senet, P.; Liwo, A.; Scheraga, H. A. Investigation of protein folding by coarse-grained molecular dynamics with the UNRES force field. *J. Phys. Chem. A* **2010**, *114*, 4471–4485.
- (59) Palleros, D.; Reid, K. L.; McCarty, J. S.; Walker, G. C.; Fink, A. L. DnaK, Hsp73, and their molten globules. Two different ways heat shock proteins respond to heat. *J. Biol. Chem.* **1992**, *267*, 5279–5285.
- (60) Ha, J.-H.; McKay, D. B. Kinetics of nucleotide-induced changes in the tryptophan fluorescence of the molecular chaperone Hsc70 and its subfragments suggest the ATP-induced conformational change follows initial ATP binding. *Biochem.* **1995**, *34*, 11635–44.
- (61) Liberek, K.; Marszalek, J.; Ang, D.; Georgopoulos, C.; Zylicz, M. *Escherichia coli* DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2874–2878.
- (62) Revington, M.; Zhang, Y.; Yip, G. N. B.; Kurochkin, A.; Zuiderweg, E. R. P. NMR investigations of allosteric processes in a two domain *thermus thermophilus* Hsp70 molecular chaperone. *J. Mol. Biol.* **2005**, *349*, 163–183.
- (63) Ma, J.; Karplus, M. Ligand-induced conformational changes in ras p21, a normal mode and energy minimization analysis. *J. Mol. Biol.* **1997**, *274*, 14–131.
- (64) Cui, Q.; Li, G.; Ma, J.; Karplus, M. A normal mode analysis of structural plasticity in the biomolecular motor F1-ATPase. *J. Mol. Biol.* **2004**, *340*, 345–372.
- (65) Gaillard, T.; Martin, E.; San Sebastian, E.; Cossio, F. P.; Lopez, X.; Dejaegere, A.; Stote, R. H. Comparative normal mode analysis of LFA-1 integrin I-domains. *J. Mol. Biol.* **2007**, *374*, 231–249.
- (66) Cecchini, M.; Houdusse, A.; Karplus, M. Allosteric communication in myosin V: from small conformational changes to large directed movements. *PLOS Comput. Biol.* **2008**, *4* (8), e1000129.
- (67) Kong, Y.; Ma, J.; Karplus, M.; Lipscomb, W. N. The allosteric mechanism of a yeast chorismate mutase, a dynamic analysis. *J. Mol. Biol.* **2006**, *356*, 237–247.

- (68) Ma, J.; Karplus, M. The allosteric mechanism of the chaperone GroEL: a dynamic analysis. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8502–8507.
- (69) Woo, H. J.; Jiang, J.; Lafer, E. M.; Sousa, R. ATP-induced conformational changes in Hsp70: molecular dynamics and experimental validation of a silico predicted conformation. *Biochem.* **2009**, *48*, 11470–11477.
- (70) Taneva, S. G.; Moro, F.; Velazquez-Campoy, A.; Muga, A. Energetics of nucleotide-induced DnaK conformational states. *Biochemistry* **2010**, *49*, 1338–1345.
- (71) Bhattacharya, A.; Kurochkin, A. V.; Yip, G. N. B.; Zhang, Y.; Bertelsen, E. B.; Zuiderweg, E. R. P. Allostery in Hsp70 chaperones is transduced by subdomain rotations. *J. Mol. Biol.* **2009**, *388*, 475–490.
- (72) Zhang, Y.; Zuiderweg, E. R. P. The 70-kDa heat shock protein chaperone nucleotide-binding domain in solution unveiled as a molecular machine that can reorient its functional subdomains. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10272–10277.

CT1002169