Received: 30 March 2022 | Revised: 6 August 2022 | Accepted: 13 September 2022

DOI: 10.1002/prot.26425

RESEARCH ARTICLE

PROTEINS WILEY

Comparison of allosteric signaling in DnaK and BiP using mutual information between simulated residue conformations

Markus Schneider | Iris Antes†

TUM Center for Functional Protein Assemblies and TUM School of Life Sciences, Technische Universität München, Freising, Bavaria, Germany

Correspondence
Markus Schneider, TUM Center for Functional Protein Assemblies and TUM School of Life Sciences, Technische Universität München, Klebelstrasse 5, 85356 Freising, Bavaria, Germany.
Email: markusg.schneider@tum.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Numbers: SFB749/C08, SFB 1035/A10

Abstract

The heat shock protein 70 kDa (Hsp70) chaperone system serves as a critical component of protein quality control across a wide range of prokaryotic and eukaryotic organisms. Divergent evolution and specialization to particular organelles have produced numerous Hsp70 variants which share similarities in structure and general function, but differ substantially in regulatory aspects, including conformational dynamics and activity modulation by cochaperones. The human Hsp70 variant BiP (also known as GRP78 or HSPA5) is of therapeutic interest in the context of cancer, neurodegenerative diseases, and viral infection, including for treatment of the pandemic virus SARS-CoV-2. Due to the complex conformational rearrangements and high sequential variance within the Hsp70 protein family, it is in many cases poorly understood which amino acid mutations are responsible for biochemical differences between protein variants. In this study, we predicted residues associated with conformational regulation of human BiP and *Escherichia coli* DnaK. Based on protein structure networks obtained from molecular dynamics simulations, we analyzed the shared information between interaction timelines to highlight residue positions with strong conformational coupling to their environment. Our predictions, which focus on the binding processes of the chaperone's substrate and cochaperones, indicate residues filling potential signaling roles specific to either DnaK or BiP. By combining predictions of individual residues into conformationally coupled chains connecting ligand binding sites, we predict a BiP specific secondary signaling pathway associated with substrate binding. Our study sheds light on mechanistic differences in signaling and regulation between Hsp70 variants, which provide insights relevant to therapeutic applications of these proteins.

KEYWORDS
allostery, HSP70 heat-shock proteins, molecular dynamics simulation, protein structure networks, signal transduction

† Died August 4, 2021.

1 | INTRODUCTION

The heat shock protein 70 kDa (Hsp70) molecular chaperone is a class of proteins found across a wide range of prokaryotic and eukaryotic organisms, with no fewer than 13 isoforms in humans alone.[1–3] Their functional roles include protein (re-)folding, membrane translocation, regulation of apoptosis, and disaggregation of denatured proteins in cooperation with other chaperone systems.[1–7] Being both ubiquitous and critical to cell damage mitigation, they are also of therapeutic interest for a variety of conditions such as cancer and neurodegenerative disorders.[2,3,7–12] The ATP-driven conformational cycle allows Hsp70s to transiently bind exposed hydrophobic stretches of substrate proteins and selectively modulate their folding process. Structurally, a Hsp70 protein is divided into a number of distinct modular domains: The nucleotide-binding domain (NBD) is an actin-like ATPase with two rotatable lobes (NBD-I and NBD-II), connected to the substrate binding domain (SBD) via a flexible linker region (NBD–SBD linker). The SBD is further divided into a β-sandwich core forming a cleft for binding of target peptides or proteins (SBDβ) and an α-helical lid which can dynamically open and close over the binding cleft (SBDα), followed by an unstructured C-terminal tail.[2,6,7,11,13,14] In the ATP-bound state of the conformational cycle, the NBD and SBDβ are predominantly docked onto each other, with the SBDα lid in the "open" conformation and stabilized by contacts with the NBD (Figure 1). In this docked conformation, the NBD–SBD linker is nestled in a cleft formed on the surface of NBD-II. Upon binding of a substrate polypeptide in the SBDβ binding cleft, the NBD–SBD interface partially undocks, allowing the NBD lobes to rotate into a position activating ATP hydrolysis. This leads to complete undocking of the NBD–SBD domains, freeing the NBD–SBD linker and stabilizing Hsp70 in a conformation with two separate domains, the NBD binding ADP and the SBD binding the substrate. The cycle is finally

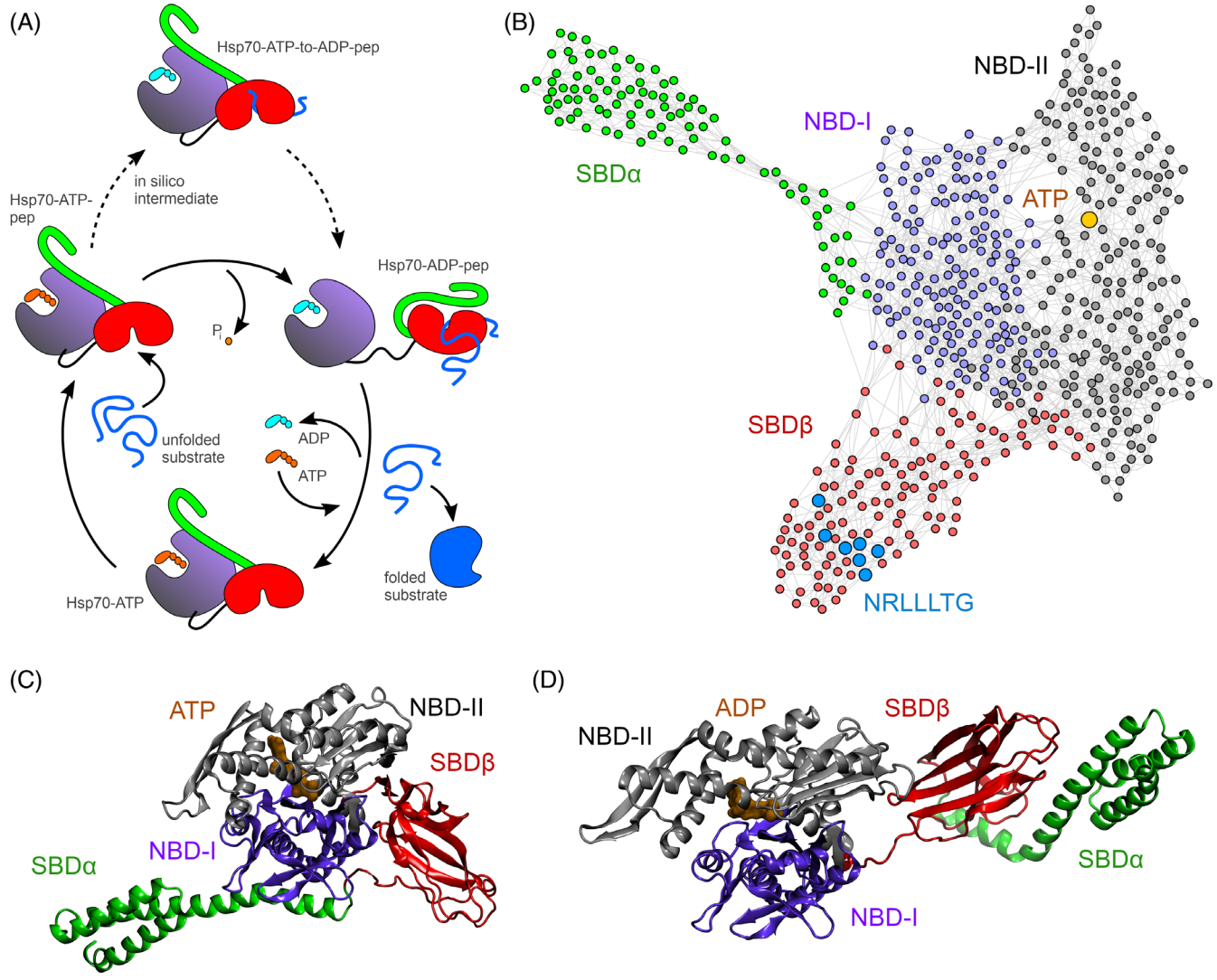

FIGURE 1 Structural organization and conformations of heat shock protein 70 kDa (Hsp70) chaperones. (A) Simplified representation of the Hsp70 conformational cycle. (B) Protein structure network of DnaK-ATP bound to the NRLLLTG peptide. (C,D) Representative structures of DnaK-ATP (C) and DnaK-ADP (D) extracted from molecular dynamics simulations with subdomain coloring. ATP/ADP nucleotides are shown in brown. NBD, nucleotide-binding domain; SBD, substrate binding domain.

completed by nucleotide exchange of ADP to ATP and subsequent re-docking of the NBD and SBD domains.[2,6,7,11,14–17] This complex orchestration is achieved through multiple pathways transmitting conformational changes and fluctuations throughout the protein, with the nucleotide and protein substrate ligands acting in concert to advance the conformational cycle.[2,6,7,11,13,14,16,18,19] In addition, co-chaperones like J-domain proteins (JDPs) or nucleotide exchange factors (NEF) accelerate these processes by forming transient complexes with Hsp70 and modulating its functional and conformational cycle.[2,6,7,11,15,20–23]

Of the diverse family of Hsp70 proteins, the *Escherichia coli* variant DnaK is by far the most extensively studied. A comprehensive series of biochemical experiments has investigated effects of point mutations on protein activity,[13] gradually assembling a model for understanding the underlying allosteric mechanisms advancing the conformational cycle. While Hsp70s from other organisms generally share the same structural architecture, it has been noted that they can differ substantially in substrate recognition, allosteric signaling, and co-chaperone interactions.[1,2,7,15,16,24–29] The human Hsp70 isoform binding immunoglobulin protein (BiP; also known as GRP78 or HSPA5) is usually found in the endoplasmatic reticulum, where it folds membrane and secretes proteins.[30,31] However, under specific conditions BiP or its isoforms can appear in certain cell types within the cytosol, nucleus, mitochondria, and on the cell surface (csBiP).[32–34] Tumors and cells under stress show elevated levels of csBiP compared to normal cells, which has prompted investigations into applications for cancer detection and therapy.[35–39] Moreover, csBiP is implicated as a coreceptor promoting host cell entry for several coronaviruses, including the pandemic virus SARS-CoV-2,[40–44] and has been suggested as a therapeutic factor for severe COVID-19 cases.[41,44] Elucidating the evolutionary differences distinguishing Hsp70 variants could deepen our understanding of this important protein class and help to tailor drugs to the specific properties of the targeted protein. The sequence homology between BiP and DnaK is below 50%, while several functional differences have been reported between the two variants, such as propensity of NBD–SBD docking, post-translational modifications or interactions with cochaperones modulating allosteric signaling.[7,16,24,31,45–47] It is mostly unknown which mutations are responsible for the observed functional and regulatory differences,[48] which are difficult to pinpoint due to the large size of Hsp70s (>600 amino acids), high sequential variance within the family and the complexity of the conformational cycle. In this work, we investigated residues associated with conformational control in the Hsp70 proteins DnaK and BiP using our recently developed difference node correlation factor (DNCF) method, which is based on estimating the shared information between interaction timelines obtained from molecular dynamics (MD) simulations and evaluating how this shared information is modulated by, for example, binding of a ligand.[49] This method implements a new variant within the class of graph-based allosteric prediction frameworks, which have been applied with success to a diverse range of proteins such as CFTR,[50] GPCRs,[51] myoglobins,[51] Hsp90,[52] and others.

Although there have been a number of studies reporting computational predictions of allostery in Hsp70s,[53–60] there is limited information on how evolutionary differences affect transferability of allosteric models between DnaK and BiP. Our predictions complement the set of experimentally determined individual residues by predicting pathways of conformationally coupled residues, which we presume to be involved in the initiation of conformational changes and allostery following a ligand binding event. On this basis, we suggest a number of residue positions which might explain the functional differences between DnaK and BiP.

2 | MATERIALS AND METHODS

2.1 | Protein structures

Structures for full length DnaK-ATP (PDB-ID: 4B9Q), BiP-ATP (PDB-ID: 5E84), and DnaK-ADP (PDB-ID: 2KHO; first model in file was used) were acquired from the RCSB PDB web site. Protein models were adjusted to reflect the sequences in Text S1 using IRECS[61] to mutate side chains in 4B9Q and MODELLER (v 9.18)[62] to mutate and add missing residues in 5E84, selecting the model with the best DOPE score out of 100 candidates. Structures containing only the NBD were derived from the full-length structures by cutting at the N-terminal side of the NBD–SBD linker. For DnaK-NBD-ADP, the structure of full-length DnaK-ADP was cut at the linker and ADP/Mg^{2+} was added based on the 4B9Q structure. The structure for BiP–NBD–ADP was created as a homology model using MOD based on a template of yeast BiP (PDB-ID: 3QFU), using the same procedure as described above. For all systems, ATP-to-ADP variants were obtained by cutting the terminal ATP phosphate from the corresponding ATP-bound structures. Ions present in crystal structures were removed with the exception of magnesium located in the nucleotide binding pocket. The NRLLLTG peptide was added to relevant systems based on the conformation found in the DnaK–peptide complex (PDB-ID: 1DKX). To allow for easier comparison between systems, indices describing residue positions were adjusted in all systems to reflect the DnaK sequence (UniProt-ID: P0A6Y8), adding PDB residue insertion codes as needed (see full sequence alignment in Text S2). Furthermore, a complete mapping of the DnaK residue indices to the UniProt sequence numbering of BiP (UniProt-ID: P11021) is provided in Table S1. Protein regions were defined by the following residue index ranges, based on DnaK: NBD-I from 1 to 177, NBD-II from 178 to 383, SBDβ from 384 to 506, and SBDα from 507 to 603.

2.2 | MD simulations

MD simulations were performed using the Amber16-AmberTools16/17 software suite[63] with the Amber14SB force field,[64] and TIP3P water[65] using ATP/ADP parameters from Meagher et al.[66] The system was solvated in a cubic water box using a minimum solute-face distance of 12 Å and neutralized with NaCl. For

the nonbonded interactions a 12 Å direct space cutoff and particle mesh Ewald (PME) summation for long-ranged electrostatic interactions were applied. Energy minimization was performed until convergence to 0.01 kcal mol^{-1} $Å^{-1}$ was reached using the XMIN minimizer. Afterwards, the volume of the solvent box was adjusted to a density of 1.00 kg m^3. Systems were gradually heated from 0 to 300 K over 1.5 ns using a variant of our published heatup protocol,[67] restraining all heavy atoms with a force constant of 3.00 kcal mol^{-1} $Å^{-2}$ until 20 K and all protein backbone atoms until 200 K. SHAKE[68] was applied to all bonds involving hydrogen and an integration time step of 1 fs was used during heatup, increasing to 2 fs for subsequent production runs. For heating and temperature control, a Langevin thermostat was used with a collision frequency of 4 ps^{-1}, and beginning from the final 0.5 ns of the heatup, a Berendsen barostat was employed with a relaxation time of 1 ps. For each system, three independent replica runs were simulated for 400 ns each, starting from separate heatup runs and with randomized Langevin seeds. The initial 100 ns of each run were removed before analysis to reduce bias toward initial structures. Atom interactions were extracted from MD trajectories with CPPTRAJ,[69] using the "nativecontacts" command for contact timelines (distance cutoff 5 Å; saving both native and nonnative time series), and the "hbond" command for hydrogen bonds (distance cutoff 3.5 Å; angle cutoff 135°).

2.3 | Protein structure networks

For analyses of protein structure networks and related quantities we used our network analysis tool SenseNet (version 1.1.0),[49] a plugin for Cytocape 3 (version 3.6.1).[70] CPPTRAJ outputs of contact and hydrogen bond timelines were processed using AIFgen[49] and loaded into SenseNet. Edges representing interactions occurring in less than 10% of the total simulation time were removed from the networks to minimize the influence of spurious interactions. DNCF scores were calculated in SenseNet as described before,[49] using the "Correlation" function set to the "Neighbor" and "Mutual information difference" modes. The obtained edge scores were then summed up using the "Degree" function. Edges of the two networks were considered equivalent if they connected the same residues and were of the same interaction type (edge mapping set to "Match Location"). As reference for DNCF calculations, we selected the corresponding networks from Hsp70-ATP (for analyses of the full-length protein) or Hsp70-NBD-ATP (for analyses of the isolated NBD domain). The DNCF method evaluates the changes in conformational coupling in neighboring interactions between a target and a reference simulation, for example, between Hsp70-ATP and Hsp70-ATP-to-ADP. Contacts and hydrogen bonds between residues are described as a timeline encoding the number of interactions in each time frame of the MD trajectory. The DNCF score is calculated as follows: Each carbon–carbon contact and each hydrogen bond in the network is represented by a separate edge X in the network. Another edge Y is said to be neighboring if it shares at least one node with X. In other words, the neighboring interactions represented by X and Y share at least one common residue (e.g., two different hydrogen bonds formed by one residue to different interaction partners). For each pair of neighboring interactions X and Y in the target simulation, the equivalent interactions $\widehat{X}$ and $\widehat{Y}$ are obtained from the reference simulation. Then, the change in shared information of the selected interaction pair between timelines from the target and reference simulation is evaluated using the difference in pointwise mutual information as

$$I(X;Y) = \sum_{x\in\cup\left(x,\widehat{x}\right)}\sum_{y\in\cup\left(y,\widehat{y}\right)}\left|p(x,y)\cdot\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)-\widehat{p}(x,y)\cdot\log_2\left(\frac{\widehat{p}(x,y)}{\widehat{p}(x)\widehat{p}(y)}\right)\right|, \quad (1)$$

with $\widehat{X},\widehat{Y}$ denoting the timelines from the reference simulation matching the locations of X and Y of the target simulation, and $p,\widehat{p}$ representing the probabilities of interaction states within the target and reference timelines. Finally, the DNCF score for each residue is obtained by summing the contributions of Equation (1) for all interactions that residue is participating in. More methodological details, including an extensive discussion on network parameters and simulation setups, can be found in our previous work.[49]

Random walks weighted by DNCF scores ("DNCF-RW") were performed using the "Random Walk" function of SenseNet in "Targeted Symmetric" mode, starting from the node representing the central leucine of the NRLLLTG peptide substrate and stopping the search when the ATP node was reached (or vice versa). Given a starting ("current") node for the random walk, the next node to be visited is selected from the list of connected neighbor nodes with the probability distribution

$$p(i) = \frac{\mathrm{DNCF}(i)}{\sum_{n\in N}\mathrm{DNCF}(n)}, \quad (2)$$

where the candidate node i is part of the set of neighbors N, that is, nodes connected to the current node. Revisiting nodes was permitted, but their contribution was only counted once. The search was restarted if the target node was not found after 1000 steps, and in addition with a probability of 0.1 at each step, ensuring that the search rejected pathways substantially longer than the shortest possible path (see Section 3). Shortest paths between two nodes were calculated using Dijkstra's algorithm, as implemented by the "Shortest path" function of SenseNet. Plots were generated using matplotlib (version 3.0.3)[71] with pictures of molecular structures by VMD (version 1.9.3)[72] and open-source PyMOL (Schrodinger, LLC. 2010. The PyMOL Molecular Graphics System, Version 1.8.4.0).

2.4 | Analysis of NBD lobe rotation

All clustering analyses were performed with the "cluster" command of CPPTRAJ. First, trajectories were aligned to the Cα atoms of NBD-I.

Following this alignment, root mean square deviations (RMSD) calculations of the NBD lobe rotation state were performed by calculating the RMSD values of Cα atoms in NBD-II. Next, trajectory frames were hierarchically clustered by their pairwise RMSD (using complete linkage) until two clusters remained. The centroid structures of each cluster were chosen as representatives. Rotational (screw) axes describing the relative lobe motion was calculated with CPPTRAJ in a two-step process: First, the representative cluster structures obtained before were aligned to the Cα atoms of NBD-I of a reference structure, that is, clusters of Hsp70-NBD-ATP-to-ADP trajectories were aligned to the crystal structures to Hsp70-NBD-ATP. Next, structures were aligned to NBD-II of the reference, extracting the corresponding rotational and translational matrices of that motion, from which the axes and angles of lobe rotation were subsequently calculated. An analogous procedure was applied to calculate the screw axes for all individual trajectory frames in order to yield the distribution of rotation angles during simulation.

3 | RESULTS AND DISCUSSION

We set out to predict protein residues responding to different ligand configurations in the Hsp70 proteins DnaK and BiP using the DNCF analysis,[49] which is based on evaluating the mutual information between the timelines of residue interactions in a protein structure network. First, we performed MD simulations of DnaK and BiP in different configurations, that is, bound to ADP, ATP, and the peptide substrate NRLLLTG (Table S2). The set of simulated systems includes Hsp70 in the ATP bound conformation ("Hsp70-ATP") and an in-silico modeled conformation bound to ATP and the NRLLLTG peptide ("Hsp70-ATP-pep"), which was chosen to approximate the substrate binding phase of the conformational cycle (Figure 1A). The process of substrate binding leading up to ATP hydrolysis involves an intermediate structure characterized by partial undocking of the NBD–SBD interface, which is not easily accessible to experimental methods of structure determination. A structure of DnaK-ATP-pep was reported recently, though only after the production of our simulations and our analyses had concluded, and no corresponding structure is currently available for BiP.[73] In the absence of this intermediate structure at the time of this work, we created variants based on the Hsp70-ATP structures, replacing ATP with ADP in silico ("Hsp70-ATP-to-ADP" and "Hsp70-ATP-to-ADP-pep"), which represents an artificial conformation for investigating the ability of protein residues to sense the ATP terminal phosphate. Each system was simulated for a total length of 1.2 μs, distributed over three independent runs of 400 ns each. The trajectories appear stable within expected variance as observed from the evolution of RMSD and force field energy terms (Figures S1–S6). A detailed discussion of these analyses can be found in Text S3.

From these trajectories, interaction timelines of carbon contacts and hydrogen bonds were extracted, transformed into networks and subsequently analyzed with SenseNet as described in our previous publication.[49] In these networks, residues are represented as nodes, which are connected by edges corresponding to residue–residue interactions. A residue pair can be connected by either a carbon contact interaction, a hydrogen bond interaction or both (using two separate edges). Furthermore, each interaction is associated with a timeline encoding the interactions state, that is, number of interactions, between two residues at different snapshots of the simulation. For example, a timeline of "1023" associated with a hydrogen bond between residues A and B indicates the presence of 1, 0, 2, and 3 hydrogen bonds at different timeslots of the simulation. Conformational correlation can then be measured by analyzing the correlation between timelines of different interactions. In the DNCF analysis,[49] this correlation is modeled by evaluating the mutual information between the timelines of neighboring interactions in the network (see Section 2). The DNCF score can be intuitively understood as the answer to the following question: Provided that we observe 0/1/2/3 or any larger number of hydrogen bonds (or carbon contacts) between residues A and B at a particular time frame of the simulation; does this influence the likelihood of observing a specific number of hydrogen bonds (or carbon contacts) in its close environment? If there is a correlation, we quantify the amount of shared information between interaction timelines. In the final step, the DNCF score calculates by how much this shared information between timelines (of its interactions) changes between two simulations. For example, the DNCF score of DnaK-ATP-pep (with DnaK-ATP as reference) indicates whether the shared information between specific interaction timelines changes due to the introduction of the peptide ligand. The DNCF score of residue A thus corresponds to the difference of shared information (in bits) between two simulations summed over all interactions (hydrogen bonds and carbon contacts) involving residue A.

By calculating the mutual information between interaction timelines of neighboring nodes in the network, residues with strong conformational coupling to their local environment can be predicted. In this context, the term "information" should not be understood as flow of bits through the protein but as contact transfer indicating routes of correlated conformations between adjacent residues. This information is summarized into a DNCF score for each residue, which corresponds to the change in shared information encoded in the interactions between two different system configurations. It is important to note that the mutual information between residues depends not only on the residues itself but on the rigidity or packing of the environment. For example, a binding event may rigidify a protein region and in turn alter the mutual information transfer and coupling of contacts between residues. Hence, it can open new routes or pathways of mutual interaction transfer between neighboring residues. Our DNCF analysis aims at predicting residues contributing to the Hsp70 allosteric network by assessing differences in the conformational coupling of residue interactions between different Hsp70 configurations, that is, either when bound to ATP, ADP, or the NRLLLTG peptide substrate.

After performing MD simulations for the selected Hsp70 configurations, hydrophobic contacts and hydrogen bonds were extracted from the trajectories and transformed into structure networks (Figure 1B). The layout of nodes in these networks was chosen to approximate the structural organization of the protein, which allows to inspect the interfaces between subdomains and trace pathways

TABLE 1 Mann–Whitney-*U* tests for association of DNCF scores with a dataset of experimentally verified allosteric residues

System	MWU *p*-value	rocAUC
DnaK-ATP-to-ADP	1.26×10^{-11}	0.84
DnaK-ATP-pep	1.28×10^{-11}	0.84
DnaK-ATP-to-ADP-pep	4.09×10^{-13}	0.86
BiP-ATP-to-ADP	1.50×10^{-7}	0.76
BiP-ATP-pep	1.82×10^{-10}	0.82
BiP-ATP-to-ADP-pep	7.37×10^{-12}	0.84

Abbreviation: rocAUC, receiver operating characteristic.

between key regions. In the ATP bound form of DnaK, both NBD lobes form an interface to the SBD, but only NBD-I has direct contact to the SBDα lid region (Figure 1B–D). The same structural organization is observed in the network of BiP. Using Dijkstra's algorithm, we determined that depending on the system, a minimum of six to seven edges need to be traversed to reach the central leucine of the NRLLLTG peptide starting from the ATP/ADP node (Table S3). This indicates a substantial distance over which a conformational signal has to be transmitted between the nucleotide and the substrate peptide binding site located within the SBD. In order to predict which residues might respond to this process in DnaK, we first calculated DNCF scores for networks based on simulations of three different configurations, that is, DnaK-ATP-to-ADP, DnaK-ATP-pep and DnaK-ATP-to-ADP-pep. As reference configuration for the DNCF calculations, we chose the network generated from the trajectory of DnaK bound to ATP ("DnaK-ATP"). Using this setup, the DNCF scores of the DnaK-ATP-to-ADP system, using DnaK-ATP as a reference, are elevated by the conformational differences induced by the in-silico exchange of ATP to ADP. In order to evaluate the agreement of our predictions with experimental data, we compared the resulting DNCF scores (Table S4) to a set of experimentally verified residues associated with allosteric effects, as found predominantly in DnaK. This dataset is composed of residue positions for which mutations affected the coupling between binding of the polypeptide substrate, nucleotide binding at the NBD, ATP hydrolysis, and NBD–SBD docking[8,13,19,22,74–84] (Table S5). As there is no comparable dataset available for specifically "non-allosteric" residues in this system, for the purpose of evaluation we categorized all residues not present in the experimental dataset as "non-allosteric"; assuming that the majority of relevant allosteric residues are already known (in DnaK), the error induced by misallocating a presumably low number of unknown allosteric residues is expected to be limited. The vast majority of experimentally verified residues were determined in DnaK as the most frequently investigated Hsp70 representative, whereas available data for other Hsp70 variants was too limited to allow for quantitative validation. In addition, we were careful to exclude functional mutants with no clear relation to an allosteric effect, that is, a mutation that was more likely to influence ligand binding affinities than communication. We began our evaluation by observing the distribution of DNCF scores within the networks, finding that all systems diverged substantially from the hypothetical normal distribution, with a notable tendency toward a log-normal shape (Figure S7). Therefore, we used the nonparametric Mann–Whitney-*U* (MWU) test to evaluate whether known allosteric residues exhibited higher DNCF scores, and found a significant ($p < .01$) increase in all tested systems (Table 1). Next, the difference between these two groups was quantified using the area under the receiver operating characteristic curve (rocAUC). The DNCF scores of all DnaK systems achieved rocAUC values of ≥ 0.84, with DnaK-ATP-to-ADP-pep yielding the top rocAUC of 0.86 (Table 1, Figure S8). Substantial association of DNCF scores with the experimental set of allosteric residues is also observed for the corresponding BiP simulations, although rocAUC scores are reduced by 0.02–0.08. Intuitively, the rocAUC indicates the probability of a randomly selected allosteric residue having a higher DNCF score than a randomly selected non-allosteric residue; the observed rocAUC decrease in BiP systems thus corresponds to a lower probability of correctly ranked residue pairs by 2%–8%. This decrease might be caused by subtle differences between the allosteric networks of BiP and DnaK, as the latter was the primary source for the experimental dataset. Thus, the rocAUC rankings do not necessarily indicate a difference in prediction quality between the system, but rather reflect the biases of the experimental dataset. Nevertheless, DNCF scores of BiP systems are still strongly correlated with known allosteric residues in DnaK, as prediction performance remains much higher than for a random model (rocAUC = 0.5). Overall, due to the consistently strong agreement of DNCF scores with experimental data in all systems, we conclude that our analysis is able to detect known allosteric residues in DnaK/BiP, which are important for the conformational coupling between the nucleotide binding region and the substrate binding region. From this basis we proceeded to predict additional candidates with potential coupling function, particularly those which may fulfill specific roles in either protein. As DnaK/BiP-ATP-to-ADP-pep consistently showed the best agreement with experimental data, we chose to focus on these configurations for further in-depth analyses.

We next investigated the structural distribution of DNCF scores within Hsp70-ATP-to-ADP-pep networks (Figure 2). The highest scoring residues, that is, within the top 10% of the network distribution, were extracted (Table 2) and mapped to the protein structures (Figure 3). Beginning with the DnaK network, Figure 2A,C shows that high-scoring residues are organized into localized clusters. The majority of residues with high DNCF scores are located in proximity to the shortest network path between the ATP/ADP nucleotide and NRLLLTG peptide (Figure 3A,B). This aligns with the experimental dataset of allosteric residues (Table S5), which were primarily determined by investigating the coupling between ATP hydrolysis and peptide binding. An additional cluster extends from the direct NBD–SBD pathway into a separate region, close to the NBD–SBD linker region and the binding site of the J-Protein cochaperone DnaJ[85] (Figures 2A, 3B). The linker itself (residues 388–394) exhibits slightly higher than average DNCF scores, less than expected considering the linker's well-established importance for controlling the NBD–SBD docking dynamics.[18,56] However, several of the adjacent high scoring residues are involved in the interface between DnaK and DnaJ (Figures 2A and

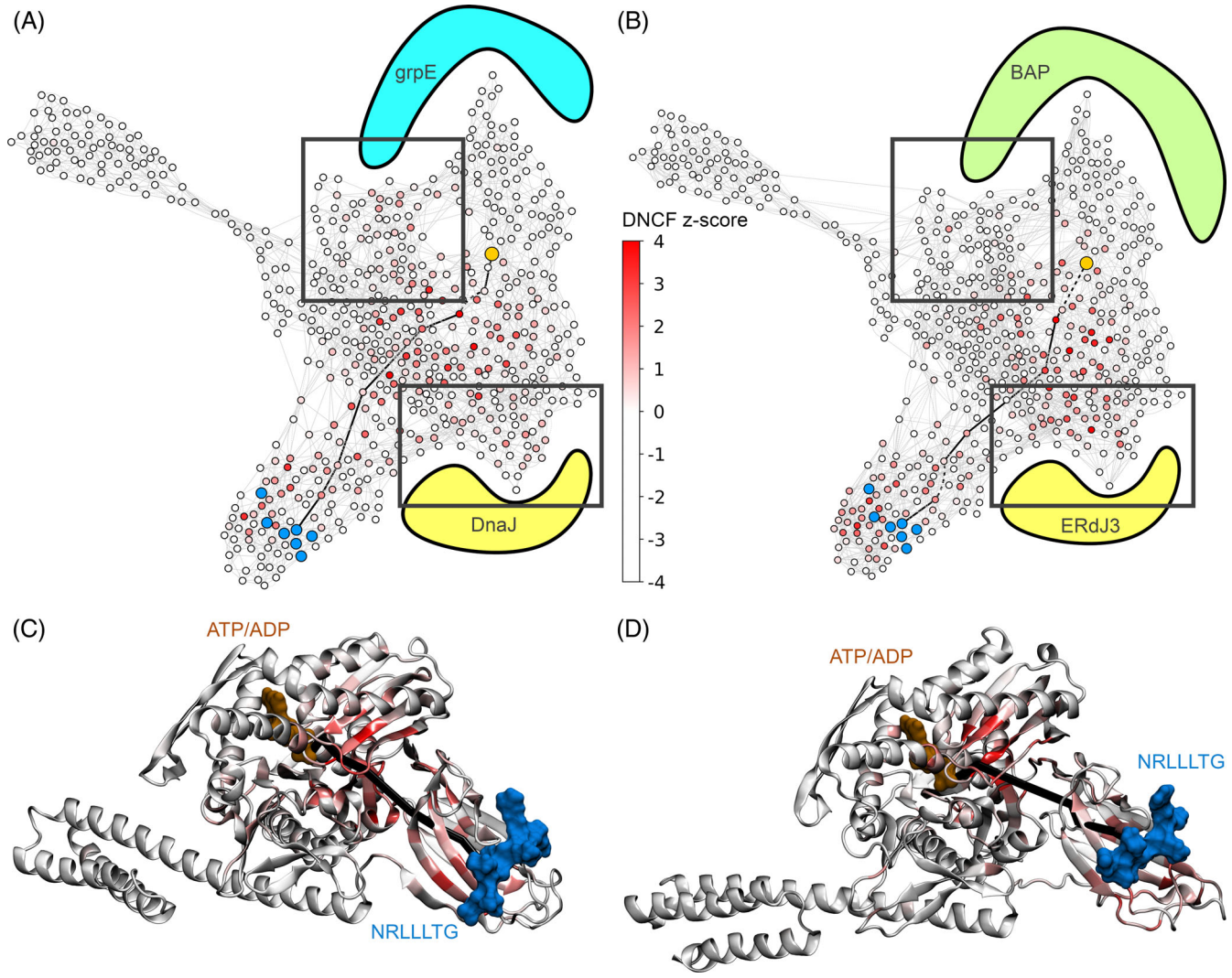

FIGURE 2 Structures and residue interaction network of heat shock protein 70 kDa proteins. (A,B) Protein structure networks obtained from molecular dynamics simulations of DnaK-ATP-to-ADP-pep (A) and BiP-ATP-to-ADP-pep (B). Nodes are colored according to the *z*-score normalized DNCF scores of their associated residues. Cochaperones DnaJ and grpE are indicated as colored shapes to visualize the location of their DnaK binding sites as observed from PDB structures (PDB-IDs: 5NRO, 1DKG), but were not present during simulations. Corresponding locations for the BiP cochaperones are estimated by homology: BAP from yeast Sil1 (PDB-ID: 3QML) and ERdJ3 from *Escherichia coli* DnaJ (PDB-ID: 5NRO). (C,D) Representative structures extracted from molecular dynamics simulations of DnaK-ATP-to-ADP-pep (C) and BiP-ATP-to-ADP-pep (D) with residues colored according to their *z*-score normalized DNCF scores

3B), which plays a substantial role in initiating the undocking of the NBD–SBD domain preceding ATP hydrolysis.[16,85] The observation that residues surrounding the linker have higher scores than the linker itself suggests that DnaJ binding may trigger a cascade of conformational changes involving residues such as R167, I168, I169, I207, K214, T395, and D481 (Figure 3B), leading to subsequent unbinding of the actual linker residues. Residues R167, I168, I169, and D481 are known to affect ATP hydrolysis and/or its stimulation by DnaJ,[19] while I207 was found to co-evolve strongly with SBD residues.[86] During the preparation of this manuscript, a structure of DnaK-ATP-pep was published in the suggested allosterically active conformation,[73] which is characterized by partial undocking of NBD and SBD domains. The conformational differences, compared to previous crystal structures of full length DnaK, were found to be concentrated in the protein region between residues 220 and 231.[73] This corroborates with a large cluster in our predictions, namely T221, N222, T225, H226, L227, and D231 (Figure 3A,B). We expect that it should be interesting to include MD simulations based on this conformation in future analyses, provided the corresponding structure can be obtained for BiP. In summary, we were able to find several localized clusters of predicted allosteric residues in DnaK, of which a substantial number are supported by previously established experimental evidence.

In addition to clusters characterized by distinct structural regions, DNCF scores also show a tendency to cluster within the protein sequence (Figure 4A). Both the SBDβ and NBD domains (including the NBD-I and NBD-II lobes) contribute high DNCF scores (Figure S9), and only the SBDα domain appears to lack any pronounced residues. Analyzing the localization of high scoring residues in more detail, we

TABLE 2 Prediction of residues which contribute to the coupling between substrate and nucleotide binding in DnaK/BiP according to DNCF scores

System	Residues
DnaK-ATP-to-ADP-pep	L9 A58 F67 **K70** R71 R75 E81 R84 T141 V142 **P143** A144 **Y145 F146 N147 D148 R151** I168 I169 **N170 E171** P172 Y193 L195 **T199** F200 **D201 I202** I207 K214 **T221** N222 T225 H226 **L227** D231 L320 V340 T395 L397 L399 I401 T437 I438 Q442 L454 Q456 F457 N458 L459 I472 F476 **D481** D490 Q497
BiP-ATP-to-ADP-pep	L9 F42 **K70** R71 R75 K100 V142 A144 **Y145** I169 **N170 E171 T173** I177 Y179 G180 L181 R184 F193 **D194** L195 **T199** F200 **D201 V202** L205 I207 F216 **T221** N222 T225 **L227** E230 F232 V340 Q384B T386 **L389 L391 D393** L399 I401 E402 M408 L411 T428 T435 V436 I438 E442 F457 L459 T460 I462 P470 Q471 I472 F476 D490 T493 N497 K498

Note: Residue positions corroborated by the set of experimentally verified residues are marked in bold font.

find particular enrichment at the subdomain interfaces between NBD-I, NBD-II, and SBDβ, while residues in the protein core and the NBD–SBDα interface trend toward lower scores (Figure S10). These observations suggest that the extensive subdomain interfaces formed in DnaK play a key role for conformational control, potentially by modulating residue packing and flexibility (e.g., NBD lobes) and changes in the equilibrium of interdomain binding (e.g., NBD–SBD docking). Comparing the DNCF score distributions between DnaK and BiP, we find that the trends for the NBD–SBD interface and its linker region are highly similar between these related proteins (Figures 2B,D, 3C,D, 4B, S11, and S12). However, while the rough structural locations of allosteric regions seemed well preserved in general, the sets of predicted residues in the top 10% percentile diverge substantially (Table 2), and in addition include residues which are unique to either protein variant, such as K214 in DnaK and Q384b in BiP. Therefore, we next focused our efforts on investigating the differences between the predicted sets of allosteric residues for DnaK and BiP.

Having observed similar DNCF score distributions between DnaK and BiP simulations, we set out to determine which residues were shared between both proteins or specific to either protein variant. The correlation between DNCF scores of DnaK-ATP-to-ADP-pep and BiP-ATP-to-ADP-pep (Figure 4C) is lower than the average between different configurations of the same protein (Spearman's *r*: 0.74 vs. average of simulations involving BiP: 0.87 ± 0.02 or DnaK: 0.9 ± 0.02) (Table 3), which suggests systematic differences between DnaK and BiP. Based on these differences, we created residue sets of predictions specific to each Hsp70 variant, that is, likely to contribute to allosteric signaling in one system but not in the other. For this, we selected residues which were specific to DnaK or BiP, that is, residues which were (i) within the top 15% of the DNCF scores in DnaK-ATP-to-ADP-pep as well as (ii) concurrently in the lower 15% of the log-normal DNCF score distribution estimated for the experimentally determined allosteric set in BiP-ATP-to-ADP-pep, and vice versa. In addition, we selected the residues which occurred in the top 10% of both systems as the "common" set of conserved allosteric residues (Table 4). The regions containing conserved allosteric residues (Figure 5) resemble the clusters of top scoring residues detected before (Figure 3). Out of the 30 allosteric candidates predicted specifically in either DnaK or BiP, 13 are related to amino acid mutations or insertions (Table 4). Residue positions with specific differences between DnaK and BiP are found in several regions: The first cluster, which is specific to DnaK, (Figure 5A,B) contains residues which contribute to the NBD–SBDβ interface (N147, D148, Q150, D481). Mutational studies have shown that these residues are important for stabilizing the NBD–SBDβ interface as well as allosteric signaling in DnaK.[19] The fact that these residue positions do not feature as prominently in BiP in our predictions suggests diminished dynamics at these locations compared to DnaK. This interpretation is backed by experimental data: Introduction of the D481N point mutation in DnaK, which is the wild type residue variant for BiP, is capable of disturbing the equilibrium of docked–undocked conformations at the NBD–SBDβ interface.[16] A similar trend toward rigidification of the same interface has been reported in multiple instances for BiP compared to DnaK.[24,31,87] In DnaK, D148 contacts the SBD via Q442 and is an essential residue for communication of the peptide binding signal from the SBD to the NBD in DnaK.[19] However, in BiP the corresponding position on the SBD side harbors a negatively charged residue (Q442E), creating electrostatic repulsion to D148. In combination, these data point toward substantial changes in the interaction pattern of the N147N-D148D-Q150Q-D481N-Q442E cluster between DnaK and BiP, which may explain differences in the dynamics of the NBD–SBD interface and allosteric communication. The second cluster is composed of residues specific to either DnaK or BiP (Figure 5B,D) and is found in the vicinity of the NBD–SBD linker (G180G, G184R, K214–, –[384b]Q, V386T, T395C), indicating another potential key region for differential regulation of NBD–SBDβ docking in these two protein variants. As it is this region that binds the J-domain in DnaK,[85] the characteristic domain shared between BiP's ERdJ1–ERdJ7 cochaperone families,[31] it appears likely that differences in the linker environment reflect evolutionary specialization to different sets of cochaperones. Another cluster, which is specific to BiP, is formed by residues located in the SBD loops (T428T, E430S, A435T, P470P, K491K, S493T, G494G, K498K, I501I). Residues 428 and 430 are part of the SBDβ's L1,2 loop and Residues 491–501 are part of the β8 sheet, two structural elements which have been shown to assume multiple distinct conformations in BiP.[88] In our data, these residues showed high DNCF scores exclusively for BiP, which might indicate increased conformational flexibility of the SBD in BiP compared to DnaK. Finally, positions 61, 62, and 65 are located close to the binding interface of NEF grpE (Figures 2A and S13) and one can speculate that these residues may be utilized to facilitate opening of the NBD loops. It is not surprising that this

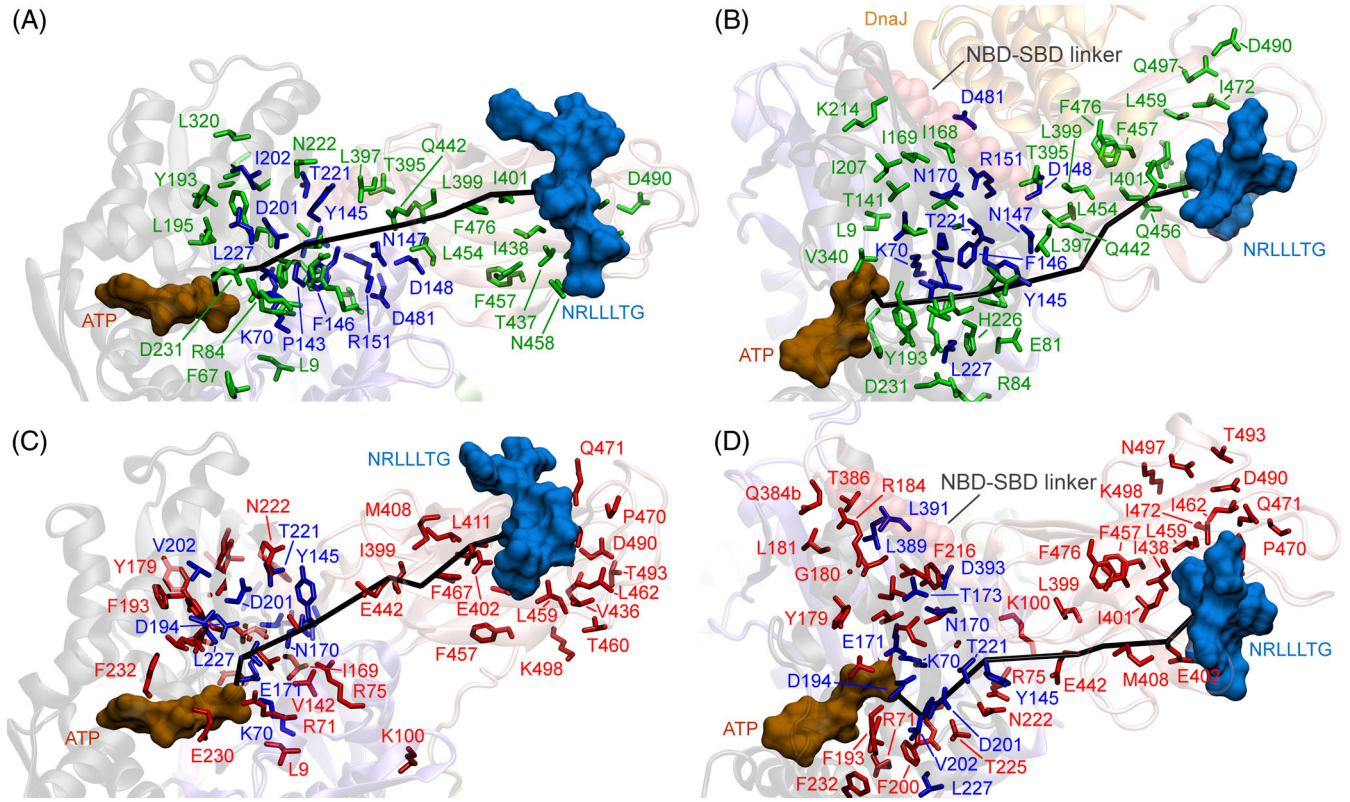

FIGURE 3 Residues showing strong conformational coupling with the nucleotide and peptide substrate according to their DNCF scores. Residues marked in blue are part of both the predicted set and the set of experimentally verified residues. The shortest inter-residue path between ADP and the central leucine of the NRLLLTG peptide is shown in black. The nucleotide-binding domain–substrate binding domain linker is indicated as red spheres. (A,B) Residues within the top 10% of DNCF scores mapped onto DnaK (blue, green). The location of the J-domain of the DnaJ cochaperone in the complex (PDB-ID: 5NRO) is indicated as an orange cartoon, though it was not present during simulation. (C,D) Residues within the top 10% of DNCF scores mapped onto BiP (blue, red)

region lacks high-scoring residues in our BiP simulations (Figure 2B), as NEFs of eukaryotic organisms including BiP (e.g., BAP[89,90]) are thought to have evolved independently from grpE and might thus utilize a different mechanism.[91]

Communication of the peptide binding signal through the protein to induce ATP hydrolysis is an essential step in the conformational cycle of Hsp70, which aligns with our observations of predicted allosteric residues clustering along the shortest path between the nucleotide binding site in the NBD and peptide binding pocket in SBDβ. However, the DNCF predictions, in isolation, account only for the conformational influence of individual residues, particularly those close to the nucleotide and substrate binding sites as these adapt fastest to the different configurations probed in our simulations, that is, in-silico exchange of ATP to ADP and the NRLLLTG peptide. Thus, conformationally coupled residues located in the intermediate region between the NBD and SBD may be overlooked within the limited ns–μs timescale of our simulations, which is much shorter than the estimated ms–s timescale characteristic for processes within the Hsp70 conformational cycle.[15,84] To address this problem, we set out to combine our predictions of individual residues into a chain of conformationally coupled residues, focusing specifically on the process of protein activation triggered by binding of the peptide substrate. We chose to perform our analyses on the Hsp70-ATP-pep systems as the configuration representing the closest approximation to that step of the conformational cycle. Starting from the node representing the central leucine of the NRLLLTG peptide in the network, we performed a weighted random walk traversing edges until the ATP node was reached, while keeping track of the visited nodes. The probability of jumping from one node to a neighboring node was chosen to be proportional to their relative DNCF scores, such that a node with twice the score than its alternative was two times as likely to be chosen for the next step (see Section 2). The procedure was then repeated after interchanging source and target nodes, that is, starting from ATP and finishing at the central leucine of the peptide. These runs, both in the forward and backward direction, were performed 10 000 times each and summed to yield the final result. This approach combines the advantages of two strategies: First, the DNCF method provides information about the conformational coupling of individual residues to their environment, and how this coupling is affected by different ligand binding states. Then, this information is supplemented with a search for the shortest paths connecting two regions, that is, nucleotide and peptide binding pockets, a technique that serves as the foundation for the class of centrality-based methods to predict functional residues in

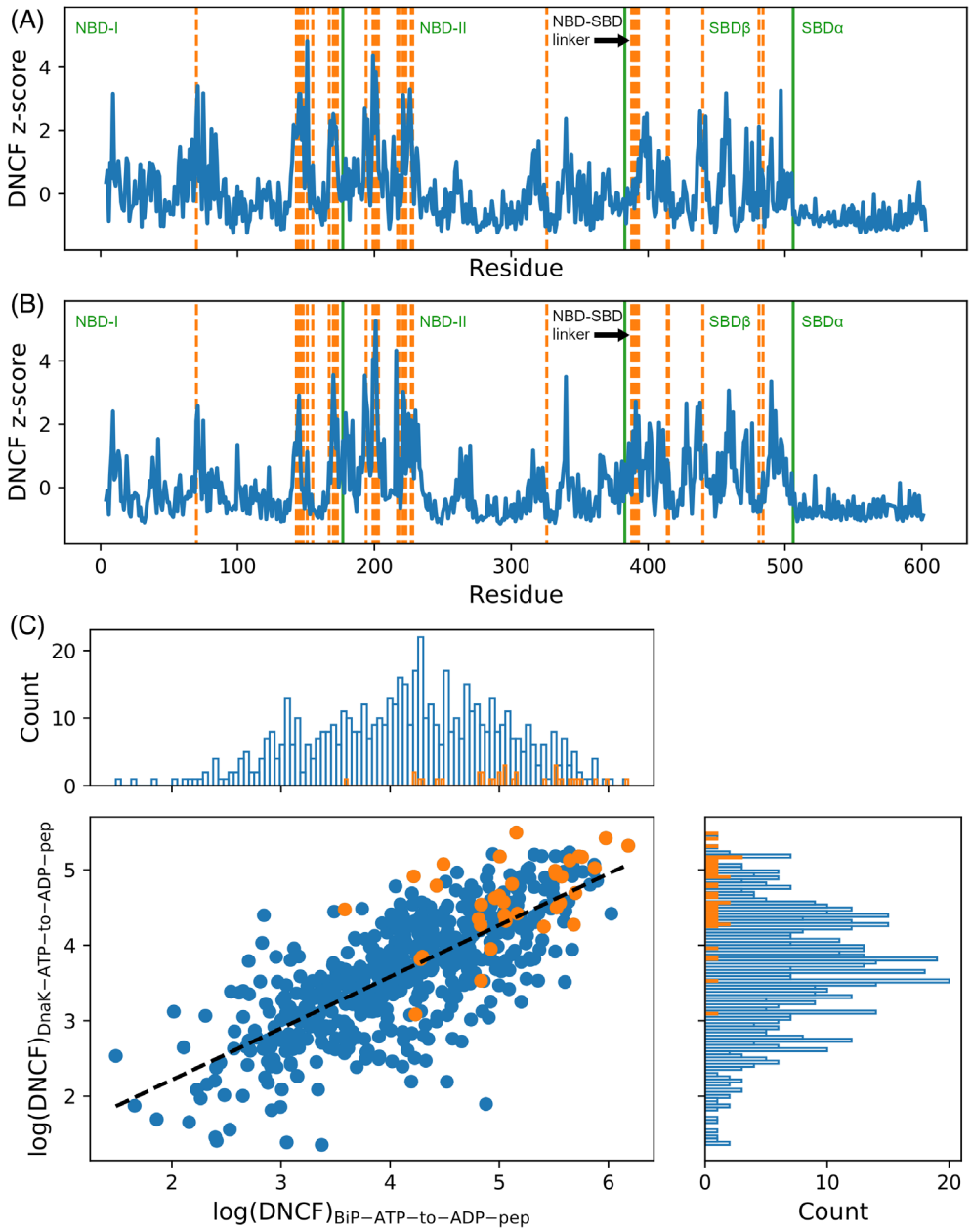

FIGURE 4 Correlation of DNCF scores obtained from molecular dynamics simulations of DnaK and BiP. Residues with experimentally verified allosteric roles are shown in orange. (A,B) Normalized DNCF scores of (A) DnaK-ATP-to-ADP-pep and (B) BiP-ATP-to-ADP-pep plotted over the protein sequence. (C) Scatterplot showing the correlation between DNCF scores of DnaK-ATP-to-ADP-pep and BiP-ATP-to-ADP-pep. NBD, nucleotide-binding domain; SBD, substrate binding domain.

TABLE 3 Spearman's correlation coefficients for DNCF scores obtained from network analysis of molecular dynamics trajectories

	DnaK-ATP-to-ADP-pep	DnaK-ATP-pep	DnaK-ATP-to-ADP	BiP-ATP-to-ADP-pep	BiP-ATP-pep
DnaK-ATP-to-ADP-pep					
DnaK-ATP-pep	0.93				
DnaK-ATP-to-ADP	0.89	0.88			
BiP-ATP-to-ADP-pep	0.74	0.72	0.61		
BiP-ATP-pep	0.76	0.74	0.65	0.89	
BiP-ATP-to-ADP	0.68	0.62	0.59	0.85	0.87

Note: System combinations featuring the same protein are highlighted in green.

proteins.[92–94] Our combined approach yields a score that accounts both for the conformational coupling of individual residues and their interconnectivity, that is, their closeness to the regions of interest and other residues with high DNCF scores. Figure 6 shows the signaling pathways predicted by this score, denoted as DNCF-RW ("DNCF Random Walk"; raw scores reported in Table S6).

TABLE 4 Prediction of residues with specific signaling properties to either DnaK or BiP according to DNCF and DNCF-RW scores

System	Method	Residues
DnaK/BiP common[a]	DNCF	L9L, K70K, R71R, R75R, V142V, A144A, Y145Y, I169I, N170N, E171E, Y193F, L195L, T199T, F200F, D201D, I202V, I207I, T221T, N222N, T225T, L227L, V340V, L399L, I401I, I438I, **Q442E,** F457F, L459L, I472I, F476F, D490D, **Q497N**
BiP[b]	DNCF	**-(106a)I**, G180G, **G184R**, E267E, K270K, **V365I**, –**(384b)Q**, **V386T**, T428T, **E430S**, **A435T**, N451N, **D460T**, P470P, K491K, **S493T**, G494G, K498K, I501I
DnaK[b]	DNCF	N61N, P62P, T65T, N147N, D148D, Q150Q, **T185E**, **K214-**, F357F, **T395C**, **D481N**
DnaK/BiP common[a]	DNCF-RW	K70K, R71R, I73I, R75R, P143P, A144A, Y145Y, F146F, N147N, D148D, Q150Q, R151R, N170N, L195L, G196G, G198G, T199T, F200F, D201D, T225T, L227L, E230E, L397L, **S398T,** L399L, I401I, E402E, M408M, L411L, F426F, V436V, T437T, I438I, V440V, L441Y, **Q442E,** L454L, **Q456T,** F457F, **N458D,** L459L, I472I, V474V, F476F, L484L
BiP[b]	DNCF-RW	P37P, **P113E**, D156D, E430S, –**(506a)R**, L507L
DnaK[b]	DNCF-RW	**R84Q**

Note: Residue codes at the beginning/end mark the DnaK/BiP sequence variants, respectively. Missing residues are indicated by a dash and insertion codes by lower case letters and parentheses. Residue positions differing between DnaK and BiP are highlighted in bold.
[a]Residues with increased scores in both systems.
[b]Residues with increased scores only in the denoted system.

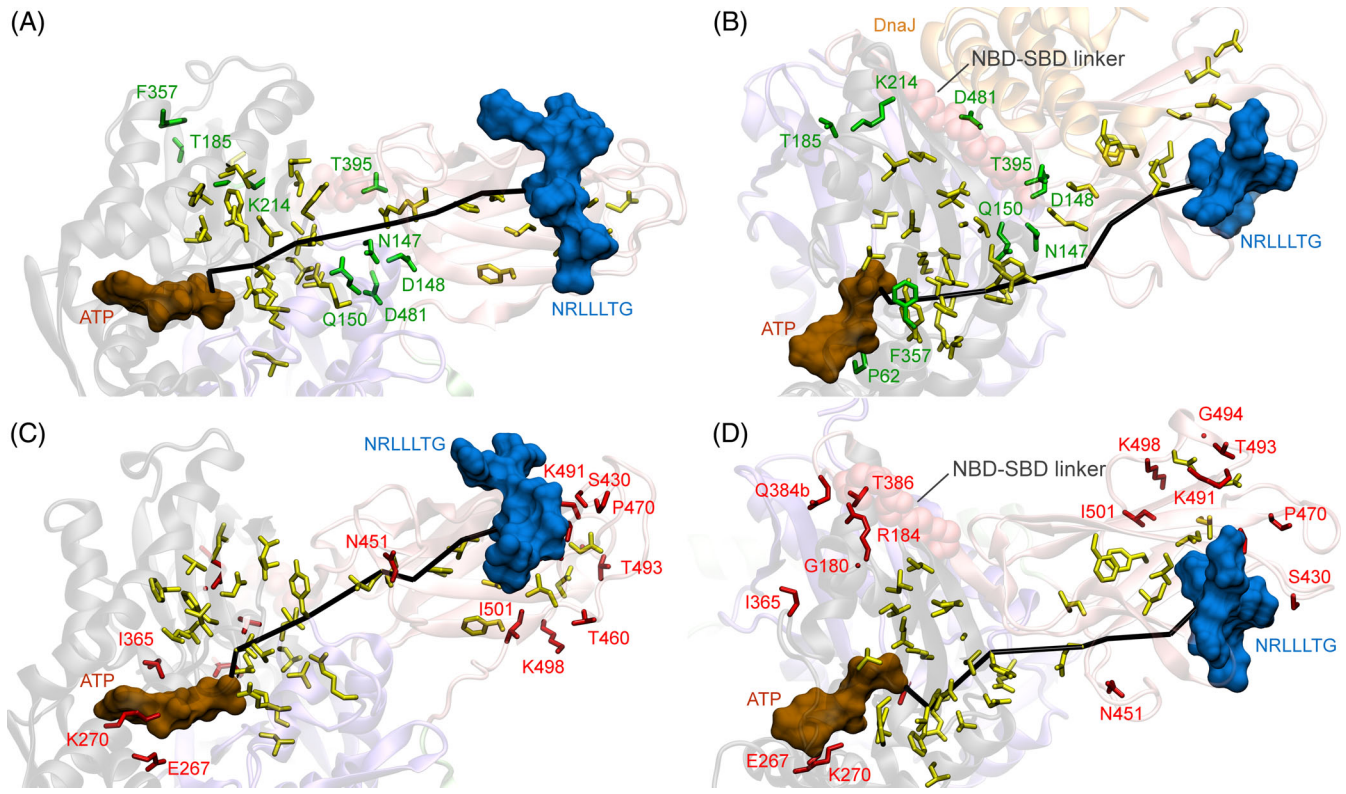

FIGURE 5 Residues showing strong conformational coupling with the nucleotide and peptide substrate specific to either DnaK or BiP as predicted by their DNCF scores. Yellow residues mark residues with increased scores in both proteins. The shortest inter-residue path between ADP and the central leucine of the NRLLLTG peptide is shown in black. The nucleotide-binding domain–substrate binding domain(NBD–SBD) linker is indicated as red spheres. (A,B) Residues with specifically increased DNCF scores in DnaK-ATP-to-ADP-pep compared to BiP (green). The location of the DnaJ cochaperone in the complex (PDB-ID: 5NRO) is indicated as an orange cartoon, though it was not present during simulation. (C,D) Residues with specifically increased DNCF scores in BiP-ATP-to-ADP-pep compared to DnaK (red)

As described above, we selected sets of shared and specific residues defined by the highest and lowest score percentiles, that is, residues which are both within the top 15% of the evaluated score in one system and within the lower 15% percentile of distribution obtained from the set of experimentally verified allosteric residues in the other system. The set of residues common to both systems consists of

TABLE 6 Predicted residues with specific roles in nucleotide-binding domain (NBD) lobe rotation compared to the full-length protein

System	Method	Residues
BiP-NBD	DNCF	V4, N64, A69
DnaK-NBD	DNCF	G229, D233, E267, K270
BiP-NBD/full length common	DNCF	L9, F42, K70, R71, R75, E171, Y179, L181, F193, D194, L195, T199, F200, D201, V202, L205, I207, F216, T225, L227, E230, F232, V340
DnaK-NBD/full length common	DNCF	L9, F67, K70, R71, R75, E81, R84, T141, Y145, Y193, L195, T199, F200, D201, I207, H226, L227, V340

45 residues, describing a contiguous surface of conformationally coupled residues between the nucleotide and substrate binding pockets (Table 4, Figure 6). R84 is the single residue found to contribute specifically to DnaK (Figure 6A,B), with the difference in scores arising from the R84Q mutation present in BiP. In contrast, the network of BiP shows a specific cluster of residues with increased scores close to the R(−506a) insertion. This residue insertion does not occur in DnaK, but is highly conserved in eukaryotes[25] and is a prominent interaction partner forming hydrogen bonds with D148, Q152, and D156 at the interface between the SBDβ core, SBDα, and NBD domains (Figure 6C,D). The R(-506a) residue has also been found to adapt to the binding of peptide substrates[28] and plays an important role in stabilizing the docking of NBD–SBD domains.[25] The cluster furthermore consists of residues P37P, P113E, D156D, and L507L,

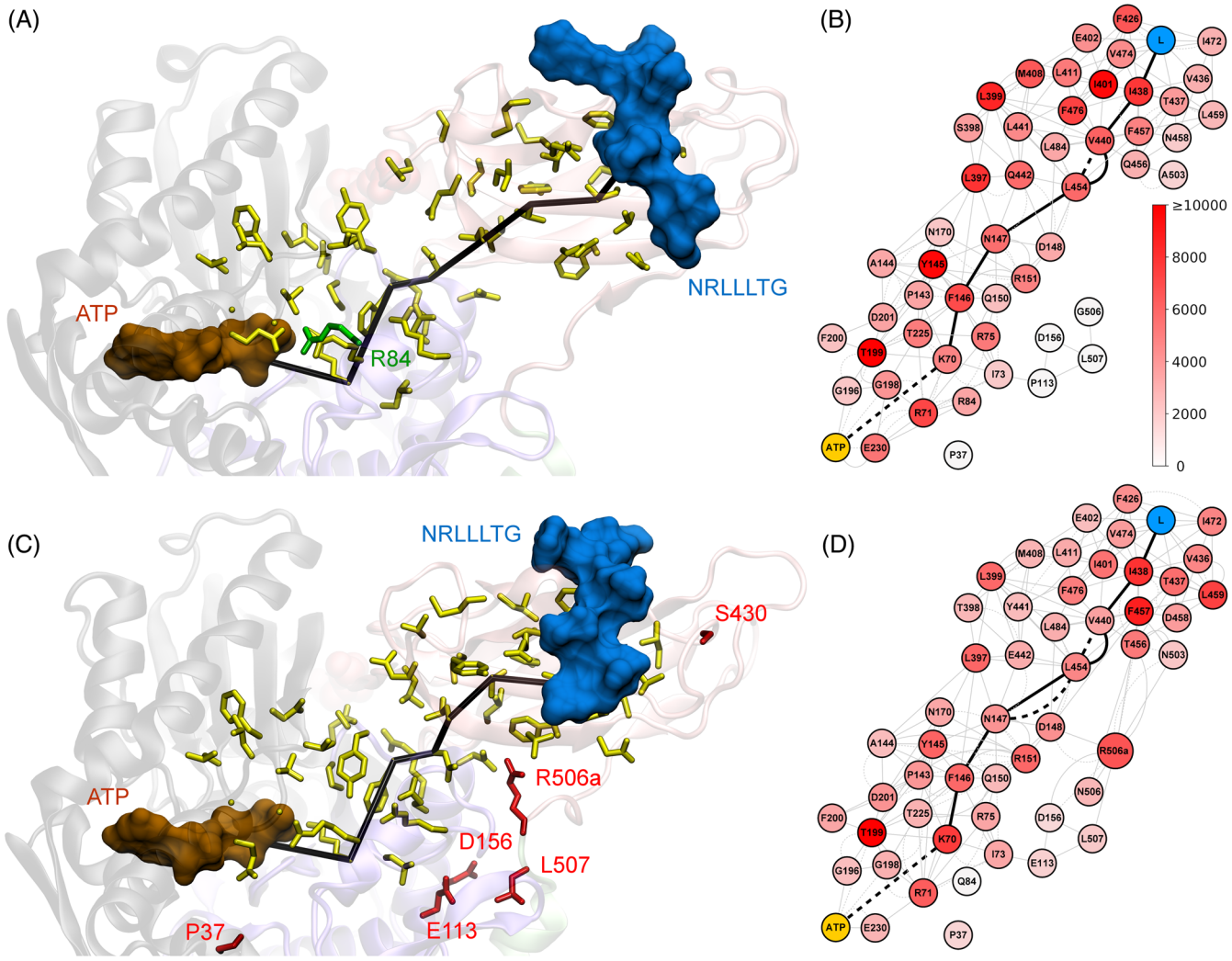

FIGURE 6 Cluster of conformationally coupled residues between the substrate and nucleotide binding sites in heat shock protein 70 kDa predicted by a targeted random walk weighted by DNCF scores (DNCF-RW). Node colors in the networks indicate the number of times each node was visited during the DNCF-RW random walk. Stick representations mark residues predicted as specific to DnaK (green), specific to BiP (red) or shared between DnaK and BiP (yellow). The shortest pathway between ADP and the NRLLLTG peptide is indicated in black. Within networks, solid edges denote carbon contacts and dashed edges indicate hydrogen bonds. (A,B) Structure and network of DnaK-ATP-pep. (C,D) Structure and network of BiP-ATP-pep

which together form an alternative pathway between NBD and SBD in addition to the main pathway present in both DnaK and BiP. In summary, our observations using the DNCF and the DNCF-RW methods suggest a high degree of conservation between DnaK and BiP with respect to the major conformationally coupled regions located between the binding pockets of the nucleotide and peptide substrate. However, differences were found in the marginal regions of the clusters, which could be the result of adaptation to their specific organism or organelle contexts. In particular, we detected three regions showing differences, which might be the result of evolutionary adaptation in DnaK and BiP: First, the NBD–SBD linker and its surrounding residues; second, the peptide binding pocket; and third, the linker region connecting the SBDβ with the SBDα Lid domain via the -(506a)R residue insert.

After investigating substrate dependent signaling in Hsp70s, we were interested in addressing the conformational effects of ATP hydrolysis on NBD lobe dynamics, particularly whether this process was controlled by the same set of residues as determined above. The partial undocking of NBD and SBD after binding a substrate peptide is the prerequisite step to prime the NBD for ATP hydrolysis. Following cleavage of the nucleotide's terminal phosphate, the NBD lobe subdomains rotate by ∼17.5°, as observed by comparing ATP-bound crystal structure and ADP- bound crystal structure of DnaK (only NBD domain; Figure S14).[84,95] In order to study the rotation of the NBD lobes, we performed simulations of the isolated DnaK/BiP NBD as an approximation to the undocked SBD state, in which both NBD and SBD are separated and only connected by the NBD–SBD linker. Three nucleotide configurations were simulated, namely Hsp70-NBD-ATP, Hsp70-NBD-ADP, and Hsp70-NBD-ATP-to-ADP. Using the RMSD of NBD-II to track the rotation state of simulations, we observed that the NBD lobes of DnaK-ATP-to-ADP rotated toward the ADP conformer in 2 out of 3 simulation replicas, while they did not rotate substantially in DnaK-ATP (Figure 7A,B). To determine whether the axis of rotation observed in the simulation matched what was expected from the crystal structures, we conducted a hierarchical clustering of trajectory frames on NBD-II until only two clusters remained. The first cluster corresponded to trajectory frames close to the initial structure, while structures showing substantial rotation to come closer to the

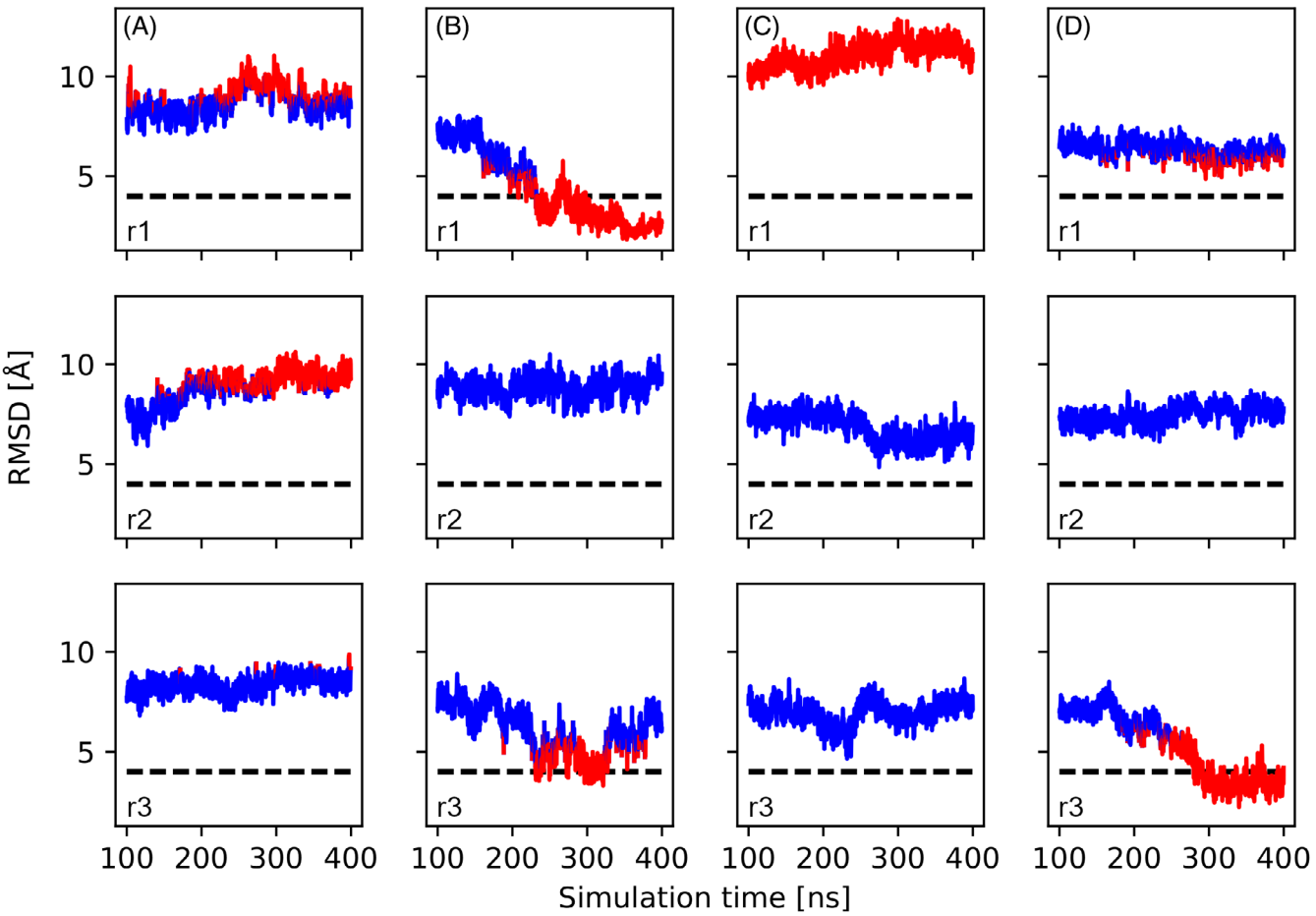

FIGURE 7 Root mean square deviations (RMSD) of nucleotide-binding domain-II (NBD-II) during molecular dynamics simulations of the isolated NBD domain of DnaK/BiP. The lobe conformation associated with bound ADP was used as reference for RMSD calculation. Trajectory frames shown in blue and red indicate membership to the two top clusters remaining after hierarchical clustering. Each column shows values obtained from replicas r1–r3 for the different systems: (A) DnaK-NBD-ATP; (B) DnaK-NBD-ATP-to-ADP; (C) BiP-NBD-ATP; (D) BiP-NBD-ATP-to-ADP

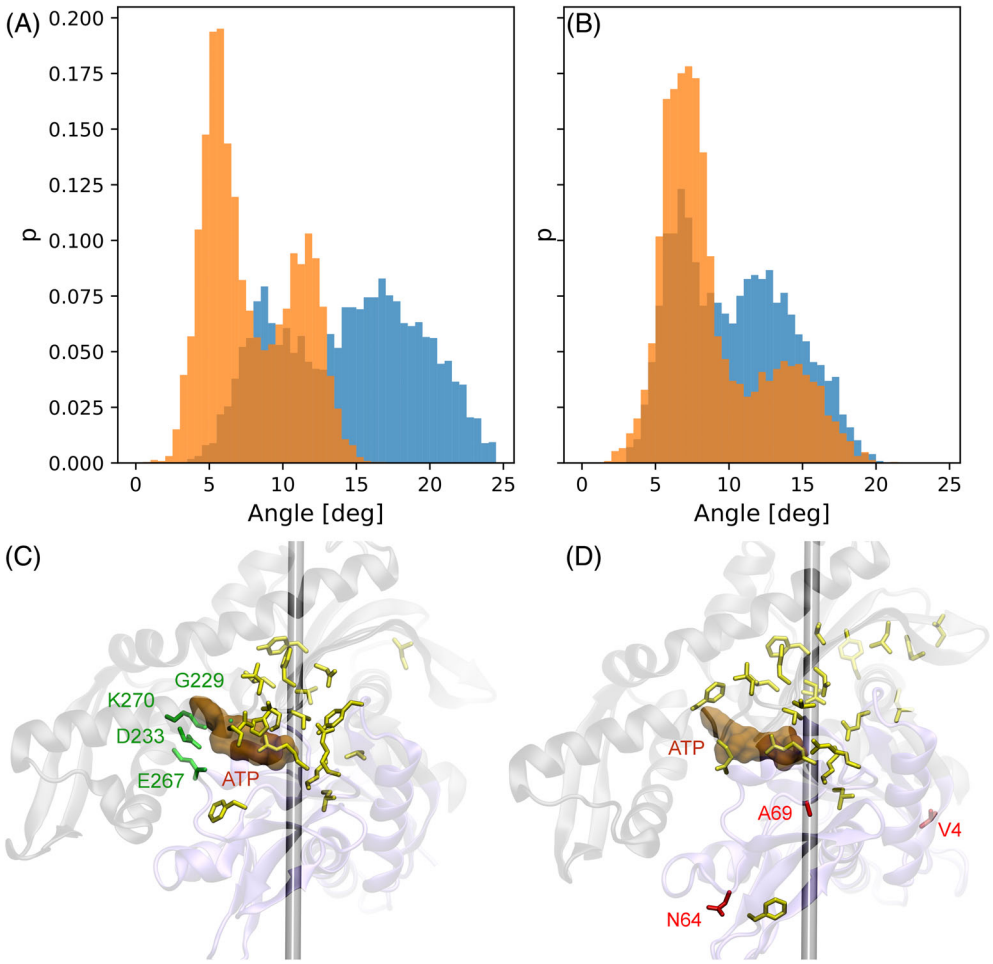

FIGURE 8 Simulation of nucleotide-binding domain (NBD) lobe rotation using molecular dynamic simulations of DnaK and BiP. (A, B) Histograms of NBD lobe rotation angles obtained from molecular dynamics trajectories of (A) BiP-NBD and (B) DnaK-NBD. The orange histograms show Hsp70-NBD-ATP while the blue histograms show the Hsp70-NBD-ATP-to-ADP variant. (C,D) Residues with specifically increased DNCF scores in the trajectories of NBD versus full length proteins in (C) DnaK (green) and (D) BiP (red)

TABLE 5 Rotation of nucleotide-binding domain (NBD) lobes during molecular dynamics simulations

System	Angle mean[a] (degree)	Angle standard deviation[a] (degree)	Angle to reference axis[b] (degree)
DnaK-NBD-ATP	4.34	2.05	78.40
DnaK-NBD-ATP-to-ADP	14.4	4.83	18.40
BiP-NBD-ATP	7.97	2.36	83.33
BiP-NBD-ATP-to-ADP	10.4	3.79	26.12

[a]Mean and standard deviation of the NBD lobe rotation angle compared to the starting conformation during molecular dynamics.
[b]Angle between the principal rotation axis of the simulation and the reference rotation axis obtained from the ATP-bound crystal structure and ADP-bound crystal structure.

DnaK-NBD-ADP conformation formed the second cluster (Figure 7A,B). The average NBD rotation during simulations, compared to the starting structure, was 14.4 ± 4.83° for DnaK-NBD-ATP-to-ADP and only 4.34 ± 2.05° for DnaK-NBD-ATP, indicating a clear increase in conformational dynamics (Figure 8A). We determined the median representative of the second cluster and calculated the rotation axis of NBD lobes compared to the initial structure as the "principal rotation" of the simulation. As reference, we calculated the rotation axis from the crystal structures of DnaK-NBD-ATP and DnaK-NBD-ADP. The principal rotation axis of DnaK-NBD-ATP-to-ADP was shifted compared to the reference axis by 18.40°, showing a much closer alignment to the reference axis than the shift of 78.4° of the DnaK-NBD-ATP system (Table 5). We performed the same analysis for BiP-NBD-ATP-to-ADP and BiP-NBD-ATP, revealing similar trends: The NBD lobes of BiP-NBD-ATP-to-ADP rotated by 10.4 ± 3.79°, while in BiP-NBD-ATP they rotated by 7.97 ± 2.36° (Figure 8B). Again, the principal rotation axis of BiP-NBD-ATP-to-ADP was much closer to the reference axis, with a shift of only 26.12°, compared to BiP-NBD-ATP with a shift of 83.33° (Table 5). This reduction in the average rotation angle of BiP compared to DnaK is explained by observing that only one out of three BiP-NBD-ATP-to-ADP replicas showed a substantial rotation toward the ADP state (Figure 7C,D). In summary, the in-silico exchange of ATP to ADP was sufficient to trigger rotation of NBD lobes toward the expected

conformers in some of our simulations. Note that we do not use this data to make a quantitative prediction of the propensity of this rotation in DnaK and BiP proteins, for which the analyzed number of replicas is too small. Instead, we utilize these simulations qualitatively to gain insight into the rough timescale on which this rotation can occur and detect candidates for residues which are specifically associated with NBD lobe rotation.

Next, we extracted interaction networks from the simulations and applied the DNCF method to determine changes in residue dynamics. Comparing the DNCF scores of DnaK-NBD-ATP-to-ADP to its corresponding full-length protein, we observe solid but not perfect correlation (Spearman's *r*: 0.82). We repeated the analyses detailed above to determine residues with specifically increased DNCF scores comparing the NBD and full-length protein simulations. Total 18 residues are found to be shared in the top 10% DNCF percentiles of DnaK-NBD-ATP-to-ADP and its full-length variant, whereas 23 residues are shared for the corresponding BiP systems (Table 6). Four residues close to the nucleotide were detected with high scores specifically in the NBD simulations, with three of them acting as direct interaction partners (G229, E267, and K270) and one located in the close vicinity (D233) (Figure 8C). This location puts them in a prime position to sense the nucleotide and provide flexibility to the NBD lobes depending on DnaK's ATP/ADP state.[58,96] In addition, G229 is located adjacent to G228, another glycine for which mutations exhibit defective chaperone function.[97] In BiP-NBD, residues V4, N64, and A69 were detected as specific for NBD lobe rotation (Figure 8D). A69 is adjacent to K70, a residue essential to ATP hydrolysis in Hsp70s,[75] indicating an association of NBD lobe rotation with conformational changes close to residues regulating catalysis. In contrast, the other two residues V4 and N64 are located far away from the axis of lobe rotation. It is possible that these observations are influenced by statistical noise, as only one out of three simulations of BiP-NBD showed NBD lobe rotation (Figure S14). Overall, the signaling properties of the isolated NBD domains appear to be very similar to the full-length protein, however with a number of residues arising with potentially specific functions for the rotational motion.

4 | CONCLUSION

In this study, we performed MD simulations of the Hsp70 chaperones DnaK and BiP, extracted networks of hydrophobic and hydrogen bond interactions and performed DNCF and DNCF-RW analyses to predict residues exchanging information about their conformational states with their environment, prompted by different ligand configurations. These residues are presumed to be associated with allosteric pathways of the Hsp70 system, that is, residues for which mutation has a notable effect on the coupling between the processes of peptide substrate binding, cochaperone-mediated activation and ATP hydrolysis. Our predictions based on the DNCF method were found to be in quantitative agreement with a set of experimentally verified allosteric residues. As the experimental dataset is limited by the number of tested mutants and reliance on DnaK as the predominant model, our predictions can aid by potentially filling gaps in our understanding of Hsp70 allostery and by pinpointing signaling differences between related protein variants, such as between DnaK and BiP. The strong agreement with experimental data further indicates that the artificial Hsp70 conformations constructed for our analysis are able to provide useful insights, despite reflecting only a part of the complete biological picture. As more and more structures of different states within the Hsp70 conformational cycle become available, like the recent publication of the partially undocked DnaK-ATP-pep conformation,[73] further MD simulations based on these novel structures will be useful for further refinement of analyses. The structures we investigated in this work—and thus the pathways we predict—correspond to one specific phase of the conformational cycle, namely the substrate mediated activation of ATP hydrolysis and subsequent undocking of the NBD–SBD interface. All simulated systems are conformationally related to the Hsp70-ATP conformation with relatively limited structural differences, that is, binding of a peptide or exchange of the nucleotide to ADP. The DNCF analysis is thus primed toward the propagation of conformational changes arising from these signal triggers. However, comparing too divergent conformations using the DNCF method, for example, the domain-docked Hsp70-ATP and the fully undocked Hsp70-ADP conformation would not yield as much useful information, as the DNCF method would simply pick up these dramatic but self-evident conformational differences. Investigating for example, the re-docking of the Hsp70-ADP conformation will require a different set of simulations, where the fully undocked Hsp70-ADP conformation is simulated alongside conformationally related variants that are more likely to initiate re-docking. Clusters formed by our predictions aligned with regions already known to be important for mediating Hsp70 conformational changes and function: The interfaces between NBD and SBD subdomains and the binding site of the JDP DnaJ. Investigated in more detail, we detected a number of residues which were predicted to be specific to either DnaK and BiP. About 40% of these differences arise directly from mutations, while others point to inherent differences between the dynamics of DnaK and BiP, such as the stability of (sub)-domain interfaces and substrate binding pocket conformational plasticity, which have been described previously on a biochemical level. By combining DNCF scores with a targeted random walk, we were able to integrate predictions of individual residues into a proposed pathway responsible for communicating binding of a substrate in the SBD to the NBD. This pathway corresponds to a series of residues whose neighboring interactions are substantially coupled and modulated by substrate and/or nucleotide binding. In this context, communication within this pathway would manifest through the possibility of subtle conformational changes or correlated fluctuations that can occur along the proposed chains of interactions. Our data revealed an alternative pathway existing in BiP but not DnaK, centered around the –(506a)R residue, which is a highly conserved position in eukaryotic Hsp70 variants.[25] Finally, we investigated the conformational control exerted by ATP/ADP over the NBD

lobes by simulating the isolated NBD domain in different configurations. We observe that the in-silico transformation of ATP to ADP is sufficient to trigger spontaneous lobe rotation during simulation toward the conformations expected from crystal structures, indicating that the terminal ATP phosphate acts as a strong mechanical wedge locking the lobes in place. Given the relatively short simulation times necessary to observe these rotations in some replicas, there appears to be a relatively low kinetic barrier to rotation after ATP hydrolysis, suggesting that this specific process does not require external assistance by cochaperones, provided that the SBDβ domain is completely undocked from the NBD. Furthermore, we identified a number of residues in DnaK which are in direct contact with the ATP/ADP nucleotide and can thus act as sensors for the nucleotide hydrolysis state, acting as focal points for initiating NBD lobe rotation. In total, our findings shed light on the pathways of allosteric communication in Hsp70s, suggesting the involvement of additional residues beyond what has been experimentally verified. We found that while many signaling residues are conserved between DnaK and BiP, there are also specific differences reflecting the divergent evolution of the two proteins. These specific residues may contribute to an explanation of the differences in biochemical behavior between Hsp70s found in different organisms and organelles. Studies elucidating differential mechanisms within a protein family provide important insights into the regulatory fine-tuning of the system, which are essential for development of targeted orthosteric or allosteric inhibitors. A possible avenue for application is indicated by a study series creating specific allosteric inhibitors for Hsp90, targeting the TRAP1 mitochondrial paralog but with no effect on cytoscolic Hsp90.[98–101] Our observations deepen our understanding of allosteric communication in the Hsp70 system and how a ubiquitous but diverse protein class has adapted to different cellular environments and cochaperone interaction partners. As Hsp70 have also been suggested to be promising therapeutic factors in a range of contexts, among them neurodegenerative diseases[2,3,7–12] and csBiP as a coreceptor of the pandemic SARS-CoV-2 virus,[40–44] investigating such evolutionary differences in further detail may become a key step in developing medical applications.

ACKNOWLEDGMENTS

We thank Martin Zacharias for constructive discussions and valuable feedback for improving the manuscript. This work is dedicated to the memory of Iris Antes, who passed away unexpectedly on the 4th of August 2021. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

This work was supported by the Deutsche Forschungsgemeinschaft (https://www.dfg.de/; SFB 1035/A10 and SFB749/C08 to Iris Antes). Markus Schneider was supported by the TUM International Graduate School of Science and Engineering (IGSSE; https://www.igsse.gs.tum.de/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CONFLICT OF INTEREST

The authors have declared that no competing interests exist.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/prot.26425.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the Dryad database (https://doi.org/10.5061/dryad.1g1jwstzs). The SenseNet Software is freely available at the Cytoscape App Store (https://cytoscape.org/) or at https://www.bioinformatics.wzw.tum.de/sensenet/method/. SenseNet is a software written and maintained by the authors, designed to be used as a plugin for the third-party network analysis tool Cytoscape 3, and distributed over the Cytoscape App Store. Cytoscape 3 can be downloaded for free at https://cytoscape.org/. The source code for SenseNet and AIFgen is included in the java archive (.jar) files used to run these programs.

ORCID

Markus Schneider https://orcid.org/0000-0002-8169-6577

REFERENCES

1. Zuiderweg ER, Hightower LE, Gestwicki JE. The remarkable multivalency of the Hsp70 chaperones. *Cell Stress Chaperones*. 2017;22(2):173-189.
2. Rosenzweig R, Nillegoda NB, Mayer MP, Bukau B. The Hsp70 chaperone network. *Nat Rev Mol Cell Biol*. 2019;20:665-680.
3. Radons J. The human HSP70 family of chaperones: where do we stand? *Cell Stress Chaperones*. 2016;21(3):379-404.
4. Balchin D, Hayer-Hartl M, Hartl FU. In vivo aspects of protein folding and quality control. *Science*. 2016;353(6294):aac4354.
5. Clerico EM, Tilitsky JM, Meng W, Gierasch LM. How hsp70 molecular machines interact with their substrates to mediate diverse physiological functions. *J Mol Biol*. 2015;427(7):1575-1588.
6. Kohler V, Andreasson C. Hsp70-mediated quality control: should I stay or should I go? *Biol Chem*. 2020;401(11):1233-1248.
7. Clerico EM, Meng W, Pozhidaeva A, Bhasne K, Petridis C, Gierasch LM. Hsp70 molecular chaperones: multifunctional allosteric holding and unfolding machines. *Biochem J*. 2019;476(11):1653-1677.
8. Zuiderweg ER, Bertelsen EB, Rousaki A, Mayer MP, Gestwicki JE, Ahmad A. Allostery in the Hsp70 chaperone proteins. *Top Curr Chem*. 2013;328:99-153.
9. Gestwicki JE, Shao H. Inhibitors and chemical probes for molecular chaperone networks. *J Biol Chem*. 2019;294(6):2151-2161.
10. Patury S, Miyata Y, Gestwicki JE. Pharmacological targeting of the Hsp70 chaperone. *Curr Top Med Chem*. 2009;9(15):1337-1351.
11. Mayer MP. The Hsp70-chaperone Machines in Bacteria. *Front Mol Biosci*. 2021;8:694012.
12. Ferraro M, D'Annessa I, Moroni E, et al. Allosteric modulators of HSP90 and HSP70: dynamics meets function through structure-based drug design. *J Med Chem*. 2019;62(1):60-87.
13. Mayer MP. Intra-molecular pathways of allosteric control in Hsp70s. *Philos Trans R Soc Lond B Biol Sci*. 2018;373(1749):20170183.
14. Mayer MP, Kityk R. Insights into the molecular mechanism of allostery in Hsp70s. *Front Mol Biosci*. 2015;2:58.
15. Mayer MP, Gierasch LM. Recent advances in the structural and mechanistic aspects of Hsp70 molecular chaperones. *J Biol Chem*. 2019;294(6):2085-2097.

16. Zhuravleva A, Clerico EM, Gierasch LM. An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell*. 2012;151(6):1296-1307.
17. Lai AL, Clerico EM, Blackburn ME, et al. Key features of an Hsp70 chaperone allosteric landscape revealed by ion-mobility native mass spectrometry and double electron-electron resonance. *J Biol Chem*. 2017;292(21):8773-8785.
18. Zhuravleva A, Gierasch LM. Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones. *Proc Natl Acad Sci USA*. 2011;108(17):6987-6992.
19. Kityk R, Vogel M, Schlecht R, Bukau B, Mayer MP. Pathways of allosteric regulation in Hsp70 chaperones. *Nat Commun*. 2015;6:8308.
20. Rosam M, Krader D, Nickels C, et al. Bap (Sil1) regulates the molecular chaperone BiP by coupling release of nucleotide and substrate. *Nat Struct Mol Biol*. 2018;25(1):90-100.
21. Karzai AW, McMacken R. A bipartite signaling mechanism involved in DnaJ-mediated activation of the *Escherichia coli* DnaK protein. *J Biol Chem*. 1996;271(19):11236-11246.
22. Laufen T, Mayer MP, Beisel C, et al. Mechanism of regulation of Hsp70 chaperones by DnaJ cochaperones. *Proc Natl Acad Sci USA*. 1999;96(10):5452-5457.
23. Marcinowski M, Höller M, Feige MJ, Baerend D, Lamb DC, Buchner J. Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. *Nat Struct & Mol Biol*. 2011;18:150-158.
24. Wieteska L, Shahidi S, Zhuravleva A. Allosteric fine-tuning of the conformational equilibrium poises the chaperone BiP for post-translational regulation. *Elife*. 2017;6:e.2943.
25. Meng W, Clerico EM, McArthur N, Gierasch LM. Allosteric landscapes of eukaryotic cytoplasmic Hsp70s are shaped by evolutionary tuning of key interfaces. *Proc Natl Acad Sci USA*. 2018;115(47):11970-11975.
26. Schneider M, Rosam M, Glaser M, et al. BiPPred: combined sequence- and structure-based prediction of peptide binding to the Hsp70 chaperone BiP. *Proteins*. 2016;84(10):1390-1407.
27. Marcinowski M, Rosam M, Seitz C, et al. Conformational selection in substrate recognition by Hsp70 chaperones. *J Mol Biol*. 2013;425(3):466-474.
28. Umehara K, Hoshikawa M, Tochio N, Tate SI. Substrate binding switches the conformation at the lynchpin site in the substrate-binding domain of human Hsp70 to enable allosteric interdomain communication. *Molecules*. 2018;23(3):528.
29. Voith von Voithenberg L, Barth A, Trauschke V, et al. Comparative analysis of the coordinated motion of Hsp70s from different organelles observed by single-molecule three-color FRET. *Proc Natl Acad Sci USA*. 2021;118(33):e2025578118.
30. Wang J, Lee J, Liem D, Ping P. HSPA5 gene encoding Hsp70 chaperone BiP in the endoplasmic reticulum. *Gene*. 2017;618:14-23.
31. Pobre KFR, Poet GJ, Hendershot LM. The endoplasmic reticulum (ER) chaperone BiP is a master regulator of ER functions: getting by with a little help from ERdj friends. *J Biol Chem*. 2019;294(6):2098-2108.
32. Lee AS. Glucose-regulated proteins in cancer: molecular mechanisms and therapeutic potential. *Nat Rev Cancer*. 2014;14(4):263-276.
33. Shin BK, Wang H, Yim AM, et al. Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function. *J Biol Chem*. 2003;278(9):7607-7616.
34. Ni M, Zhang Y, Lee AS. Beyond the endoplasmic reticulum: atypical GRP78 in cell viability, signalling and therapeutic targeting. *Biochem J*. 2011;434(2):181-188.
35. Arap MA, Lahdenranta J, Mintz PJ, et al. Cell surface expression of the stress response chaperone GRP78 enables tumor targeting by circulating ligands. *Cancer Cell*. 2004;6(3):275-284.
36. Kim Y, Lillo AM, Steiniger SCJ, et al. Targeting heat shock proteins on cancer cells: selection, characterization, and cell-penetrating properties of a peptidic GRP78 ligand. *Biochemistry*. 2006;45(31):9434-9444.
37. Liu Y, Steiniger SC, Kim Y, Kaufmann GF, Felding-Habermann B, Janda KD. Mechanistic studies of a peptidic GRP78 ligand for cancer cell-specific drug delivery. *Mol Pharm*. 2007;4(3):435-447.
38. Zhang Y, Liu R, Ni M, Gill P, Lee AS. Cell surface relocalization of the endoplasmic reticulum chaperone and unfolded protein response regulator GRP78/BiP. *J Biol Chem*. 2010;285(20):15065-15075.
39. Gopal U, Pizzo SV. Cell surface GRP78 signaling: an emerging role as a transcriptional modulator in cancer. *J Cell Physiol*. 2021;236(4):2352-2363.
40. Carlos AJ, Ha DP, Yeh DW, et al. The chaperone GRP78 is a host auxiliary factor for SARS-CoV-2 and GRP78 depleting antibody blocks viral entry and infection. *J Biol Chem*. 2021;296:100759.
41. Katopodis P, Randeva HS, Spandidos DA, Saravi S, Kyrou I, Karteris E. Host cell entry mediators implicated in the cellular tropism of SARS-CoV-2, the pathophysiology of COVID-19 and the identification of microRNAs that can modulate the expression of these mediators (review). *Int J Mol Med*. 2022;49(2):20.
42. Das JK, Roy S, Guzzi PH. Analyzing host-viral interactome of SARS-CoV-2 for identifying vulnerable host proteins during COVID-19 pathogenesis. *Infect Genet Evol*. 2021;93:104921.
43. Chu H, Chan CM, Zhang X, et al. Middle East respiratory syndrome coronavirus and bat coronavirus HKU9 both can utilize GRP78 for attachment onto host cells. *J Biol Chem*. 2018;293(30):11709-11726.
44. Shahriari-Felordi M, Alikhani HK, Hashemian SR, Hassan M, Vosough M. Mini review ATF4 and GRP78 as novel molecular targets in ER-stress modulation for critical COVID-19 patients. *Mol Biol Rep*. 2022;49:1545-1549.
45. Behnke J, Feige MJ, Hendershot LM. BiP and its nucleotide exchange factors Grp170 and Sil1: mechanisms of action and biological functions. *J Mol Biol*. 2015;427(7):1589-1608.
46. Serlidaki D, van Waarde M, Rohland L, et al. Functional diversity between HSP70 paralogs caused by variable interactions with specific co-chaperones. *J Biol Chem*. 2020;295(21):7301-7316.
47. Li H, Musayev FN, Yang J, et al. A novel and unique ATP hydrolysis to AMP by a human Hsp70 binding immunoglobin protein (BiP). *Protein Sci*. 2021;31(4):797-810.
48. Bonomo J, Welsh JP, Manthiram K, Swartz JR. Comparing the functional properties of the Hsp70 chaperones, DnaK and BiP. *Biophys Chem*. 2010;149(1–2):58-66.
49. Schneider M, Antes I. SenseNet, a tool for analysis of protein structure networks obtained from molecular dynamics simulations. *PLoS One*. 2022;17(3):e0265194.
50. Proctor EA, Kota P, Aleksandrov AA, He L, Riordan JR, Dokholyan NV. Rational coupled dynamics network manipulation rescues disease-relevant mutant cystic fibrosis transmembrane conductance regulator. *Chem Sci*. 2015;6(2):1237-1246.
51. del Sol A, Fujihashi H, Amoros D, Nussinov R. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol Syst Biol*. 2006;2006(2):2006.0019. doi:10.1038/msb4100063
52. Blacklock K, Verkhivker GM. Computational modeling of allosteric regulation in the hsp90 chaperones: a statistical ensemble analysis of protein structure networks and allosteric communications. *PLoS Comput Biol*. 2014;10(6):e1003679.
53. Penkler D, Sensoy O, Atilgan C, Tastan BO. Perturbation-response scanning reveals key residues for allosteric control in Hsp70. *J Chem Inf Model*. 2017;57(6):1359-1374.
54. Stetz G, Verkhivker GM. Computational analysis of residue interaction networks and coevolutionary relationships in the Hsp70

chaperones: a community-hopping model of allosteric regulation and communication. *PLoS Comput Biol*. 2017;13(1):e1005299.
55. Stetz G, Verkhivker GM. Dancing through life: molecular dynamics simulations and network-centric modeling of allosteric mechanisms in Hsp70 and Hsp110 chaperone proteins. *PLoS One*. 2015;10(11):e0143752.
56. English CA, Sherman W, Meng W, Gierasch LM. The Hsp70 interdomain linker is a dynamic switch that enables allosteric communication between two structured domains. *J Biol Chem*. 2017;292(36):14765-14774.
57. Chiappori F, Merelli I, Milanesi L, Colombo G, Morra G. An atomistic view of Hsp70 allosteric crosstalk: from the nucleotide to the substrate binding domain and back. *Sci Rep*. 2016;6:23474.
58. Ung PM-U, Thompson AD, Chang L, Gestwicki JE, Carlson HA. Identification of key hinge residues important for nucleotide-dependent allostery in *E. coli* Hsp70/DnaK. *PLoS Comput Biol*. 2013;9(11):e1003279.
59. Nicolaï A, Delarue P, Senet P. Decipher the mechanisms of protein conformational changes induced by nucleotide binding through free-energy landscape analysis: ATP binding to Hsp70. *PLoS Comput Biol*. 2013;9(12):e1003379.
60. Chiappori F, Merelli I, Colombo G, Milanesi L, Morra G. Molecular mechanism of allosteric communication in Hsp70 revealed by molecular dynamics simulations. *PLoS Comput Biol*. 2012;8(12):e1002844.
61. Hartmann C, Antes I, Lengauer T. IRECS: a new algorithm for the selection of most probable ensembles of side-chain conformations in protein models. *Protein Sci: Publ Protein Soc*. 2007;16(7):1294-1307.
62. Eswar N, Webb B, Marti-Renom MA, et al. Comparative protein structure modeling using Modeller. *Curr Protoc Bioinform* 2006; Chapter 5:Unit-5 6. doi:10.1002/0471250953.bi0506s15
63. Case DA, Cerutti DS, Cheatham TE, et al. *AMBER 2017*. University of California; 2017.
64. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput*. 2015;11(8):3696-3713.
65. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926-935.
66. Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. *J Comput Chem*. 2003;24(9):1016-1025.
67. Duell ER, Glaser M, Le Chapelain C, Antes I, Groll M, Huber EM. Sequential inactivation of Gliotoxin by the S-methyltransferase TmtA. *ACS Chem Biol*. 2016;11(4):1082-1089.
68. Miyamoto S, Kollman PA. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem*. 1992;13(8):952-962.
69. Roe DR, Cheatham TE 3rd. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput*. 2013;9(7):3084-3095.
70. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-2504.
71. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng*. 2007;9(3):90-95.
72. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-38.
73. Wang W, Liu Q, Liu Q, Hendrickson WA. Conformational equilibria in allosteric control of Hsp70 chaperones. *Mol Cell*. 2021;81(19):3919-3933.e7.
74. Vogel M, Bukau B, Mayer MP. Allosteric regulation of Hsp70 chaperones by a proline switch. *Mol Cell*. 2006;21(3):359-367.
75. Barthel TK, Zhang J, Walker GC. ATPase-defective derivatives of *Escherichia coli* DnaK that behave differently with respect to ATP-induced conformational change and peptide release. *J Bacteriol*. 2001;183(19):5482-5490.
76. Burkholder WF, Panagiotidis CA, Silverstein SJ, Cegielska A, Gottesman ME, Gaitanaris GA. Isolation and characterization of an *Escherichia coli* DnaK mutant with impaired ATPase activity. *J Mol Biol*. 1994;242(4):364-377.
77. Gässler CS, Buchberger A, Laufen T, et al. Mutations in the DnaK chaperone affecting interaction with the DnaJ cochaperone. *Proc Natl Acad Sci USA*. 1998;95(26):15229-15234.
78. Vogel M, Mayer MP, Bukau B. Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. *J Biol Chem*. 2006;281(50):38705-38711.
79. Suh W-C, Burkholder WF, Lu CZ, Zhao X, Gottesman ME, Gross CA. Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. *Proc Natl Acad Sci*. 1998;95(26):15223-15228.
80. Kamath-Loeb AS, Lu CZ, Suh W-C, Lonetto MA, Gross CA. Analysis of three DnaK mutant proteins suggests that progression through the ATPase cycle requires conformational changes. *J Biol Chem*. 1995;270(50):30051-30059.
81. Smock RG, Rivoire O, Russ WP, et al. An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Mol Syst Biol*. 2010;6:414.
82. Mayer MP, Laufen T, Paal K, McCarty JS, Bukau B. Investigation of the interaction between DnaK and DnaJ by surface Plasmon resonance spectroscopy. *J Mol Biol*. 1999;289(4):1131-1144.
83. Montgomery DL, Morimoto RI, Gierasch LM. Mutations in the substrate binding domain of the Escherichia coli 70 kda molecular chaperone, DnaK, which alter substrate affinity or interdomain. *J Mol Biol*. 1999;286(3):915-932.
84. Kityk R, Kopp J, Sinning I, Mayer MP. Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. *Mol Cell*. 2012;48(6):863-874.
85. Kityk R, Kopp J, Mayer MP. Molecular mechanism of J-domain-triggered ATP hydrolysis by Hsp70 chaperones. *Mol Cell*. 2018;69(2):227-237.e4.
86. General IJ, Liu Y, Blackburn ME, Mao W, Gierasch LM, Bahar I. ATPase subdomain IA is a mediator of interdomain allostery in Hsp70 molecular chaperones. *PLoS Comput Biol*. 2014;10(5):e1003624.
87. Yang J, Nune M, Zong Y, Zhou L, Liu Q. Close and allosteric opening of the polypeptide-binding site in a human Hsp70 chaperone BiP. *Structure*. 2015;23(12):2191-2203.
88. Yang J, Zong Y, Su J, et al. Conformation transitions of the polypeptide-binding pocket support an active substrate release from Hsp70s. *Nat Commun*. 2017;8(1):1201.
89. Yan M, Li J, Sha B. Structural analysis of the Sil1–Bip complex reveals the mechanism for Sil1 to function as a nucleotide-exchange factor. *Biochem J*. 2011;438(3):447-455.
90. Shomura Y, Dragovic Z, Chang HC, et al. Regulation of Hsp70 function by HspBP1: structural analysis reveals an alternate mechanism for Hsp70 nucleotide exchange. *Mol Cell*. 2005;17(3):367-379.
91. Bracher A, Verghese J. The nucleotide exchange factors of Hsp70 molecular chaperones. *Front Mol Biosci*. 2015;2:10.
92. O'Rourke KF, Gorman SD, Boehr DD. Biophysical and computational methods to analyze amino acid interaction networks in proteins. *Comput Struct Biotechnol J*. 2016;14:245-251.
93. Greene LH. Protein structure networks. *Brief Funct Genom*. 2012;11:469-478.
94. Di Paola L, Giuliani A. Protein contact network topology: a natural language for allostery. *Curr Opin Struc Biol*. 2015;31:43-48.
95. Bertelsen EB, Chang L, Gestwicki JE, Zuiderweg ER. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proc Natl Acad Sci USA*. 2009;106(21):8471-8476.

96. Liu Y, Gierasch LM, Bahar I. Role of Hsp70 ATPase domain intrinsic dynamics and sequence evolution in enabling its functional interactions with NEFs. *PLoS Comput Biol*. 2010;6(9):e1000931.
97. Chang L, Thompson AD, Ung P, Carlson HA, Gestwicki JE. Mutagenesis reveals the complex relationships between ATPase rate and the chaperone activities of Escherichia coli heat shock protein 70 (Hsp70/DnaK). *J Biol Chem*. 2010;285(28):21282-21291.
98. Moroni E, Agard DA, Colombo G. The structural asymmetry of mitochondrial Hsp90 (Trap1) determines fine tuning of functional dynamics. *J Chem Theory Comput*. 2018;14(2):1033-1044.
99. Serapian SA, Moroni E, Ferraro M, Colombo G. Atomistic simulations of the mechanisms of the poorly catalytic mitochondrial chaperone Trap1: insights into the effects of structural asymmetry on reactivity. *ACS Catalysis*. 2021;11(14):8605-8620.
100. Sanchez-Martin C, Moroni E, Ferraro M, et al. Rational Design of Allosteric and Selective Inhibitors of the molecular chaperone TRAP1. *Cell Rep*. 2020;31(3):107531.
101. Sanchez-Martin C, Menon D, Moroni E, et al. Honokiol Bis-Dichloroacetate is a selective allosteric inhibitor of the mitochondrial chaperone TRAP1. *Antioxid Redox Signal*. 2021;34(7):505-516.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Schneider M, Antes I. Comparison of allosteric signaling in DnaK and BiP using mutual information between simulated residue conformations. *Proteins*. 2023; 91(2):237-255. doi:10.1002/prot.26425