SCIENTIFIC REPORTS

OPEN

Modulation of Human Hsp90α Conformational Dynamics by Allosteric Ligand Interaction at the C-Terminal Domain

David L. Penkler & Özlem Tastan Bishop

Recent years have seen heat shock protein 90 kDa (Hsp90) attract significant interest as a viable drug target, particularly for cancer. To date, designed inhibitors that target the ATPase domain demonstrate potent anti-proliferative effects, but have failed clinical trials due to high levels of associated toxicity. To circumvent this, the focus has shifted away from the ATPase domain. One option involves modulation of the protein through allosteric activation/inhibition. Here, we propose a novel approach: we use previously obtained information via residue perturbation scanning coupled with dynamic residue network analysis to identify allosteric drug targeting sites for inhibitor docking. We probe the open conformation of human Hsp90α for druggable sites that overlap with these allosteric control elements, and identify three putative natural compound allosteric modulators: Cephalostatin 17, 20(29)-Lupene-3β-isoferulate and 3'-Bromorubrolide F. We assess the allosteric potential of these ligands by examining their effect on the conformational dynamics of the protein. We find evidence for the selective allosteric activation and inhibition of Hsp90's conformational transition toward the closed state in response to ligand binding and shed valuable insight to further the understanding of allosteric drug design and Hsp90's complex allosteric mechanism of action.

The 90 KDa heat shock protein (Hsp90) is a highly conserved molecular chaperone crucially involved in maintaining cellular homoeostasis in organisms from most kingdoms of life with the exception of archea[1]. In the cytosol, Hsp90's main biological function is the facilitation of folding, maturation, and trafficking of numerous client peptides both native and denatured[2–4]. Hsp90's diverse array of clientele implicate the chaperone in several associated biological functions and place it at the intersection of various fundamental cellular pathways, where it acts as a central hub in maintaining numerous protein interaction networks[1].

Hsp90 exists as a homodimer (Fig. 1-A), and each protomer is comprised of three well characterized domains[5–7]: an N-terminal domain (NTD) which is responsible for ATPase activity and facilitating transient inter-protomer dimerization[8]; a middle domain (M-domain) that provides a large surface area for cofactor and client binding and contributes to ATPase activation[9]; a C-terminal domain (CTD) which serves as the primary site for inter-protomer dimerization[10,11]. The NTD and M-domain are connected by a highly flexible charged linker that has been implicated in modulating chaperone function[12–15]. Hsp90's molecular function critically hinges around its ability to bind and release client peptides via a complex nucleotide dependent conformational cycle (Fig. 1-B). In a nucleotide free state, the dimer becomes highly flexible and is capable of assuming multiple conformers with a higher affinity for an open "v-like" conformation in which the M-domains of each protomer are suitably exposed for client loading[16–18]. ATP binding triggers structural rearrangements in the NTD that promote dimerization at the N-terminal, stabilizing a closed catalytically active conformation[10,19]. Transition to the closed ATPase active state is an inherently slow process recording time constants in the order of minutes[8,20,21], possibly due to energetic barriers presented by structural intermediates that may be overcome through cofactor mediation[22–25]. ATP hydrolysis and the subsequent release of ADP from the NTD initiate a conformational return to the native apo open state and client release.

Research Unit in Bioinformatics (RUBi), Department of Biochemistry and Microbiology, Rhodes University, Grahamstown, 6140, South Africa. Correspondence and requests for materials should be addressed to Ö.T.B. (email: o.tastanbishop@ru.ac.za)
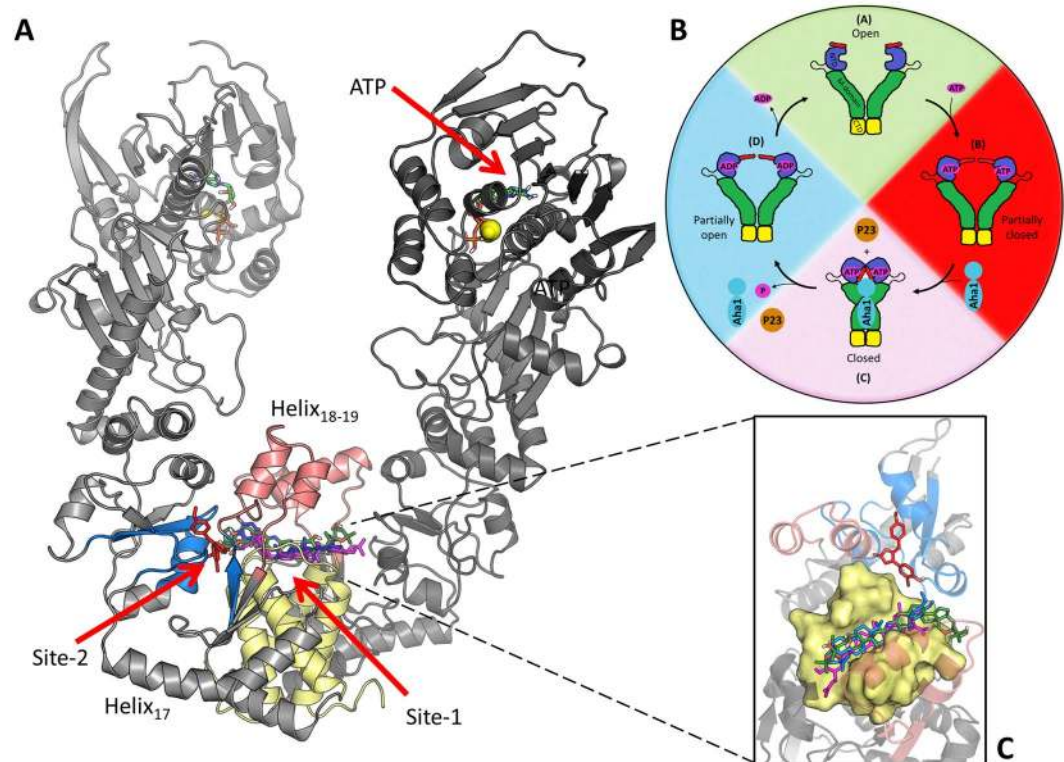
Figure 1. Illustration of Hsp90α in the open conformation. (**A**) The location of the different binding site residues are shaded: Site-1 helix$_{18\text{-}19}$ (red), helix$_{21\text{-}22}$ four-helix bundle (yellow) and Site-2 sub-pocket (blue). The NTD location of ATP and magnesium ions (spheres) are also shown. (**B**) Hsp90's nucleotide driven conformational cycle (Adopted from Penkler *et al.*[37]). (**C**) Inset – zoomed in view of docked compounds SANC309 red, SANC491 green, SANC518 blue, and Novobiocin magenta.

The interaction of Hsp90 with several disease related peptides implicate the chaperone with the progression and development of several associated pathologies such as protein folding disorders, cancer, and neurological disease[5]. In recent years it has become increasingly clear that deregulation of Hsp90 may present an attractive treatment strategy for these diseases, elevating interest in human Hsp90 as a viable drug target particularly for the treatment of cancer[26–28]. To date, Hsp90 inhibitors have largely targeted the ATPase domain, demonstrating potent anti-proliferative effects; however, their progression through clinical trials has been severely limited due to high levels of toxicity associated with an induced heat shock response[26]. For this reason, recent research has shifted focus away from the NTD to targeting distant functional sites of the protein as alternative drug treatment strategies. The design and development of inhibitors that target the formation of specific Hsp90-co-chaperone complexes, has led to the discovery of inhibitors such as Derrubone, Withaferin A, and Celastrol which block the Hsp90-Cdc37 complex[29]. Discovery of a secondary nucleotide binding site at the CTD led to the design and development of several CTD inhibitors including the coumarin based antibiotic Novobiocin that has been shown to allosterically destabilise the Hsp90 dimer by interfering with CTD dimerization leading to the dissociation of bound clients[30–32]. These CTD inhibitors of Hsp90 have a distinct advantage over the NTD ATPase inhibitors, in that they do not appear to induce secondary heat shock responses and thus incur lower levels of toxicity[29].

To date, biochemical and computational studies have demonstrated allosteric coupling between the NTD and CTD[33–37] suggesting nucleotide driven conformational restructuring may be influenced by allosteric events at the CTD. More specifically, computational studies based on the closed conformation of yeast Hsp90 have identified a putative allosteric binding site located at the M-domain:CTD inter-protomer interface that is implicated in allosteric modulation of conformational dynamics[33,35,38]. Subsequent to this discovery, further studies have led to the development of several allosteric CTD ligands that appear to enhance ATPase activity up to six-fold by promoting conformational dynamics in favour of the ATPase active closed conformation[39,40]. Furthermore, a recent biochemical study has provided evidence of Bisphenol A based CTD inhibitors of human Hsp90α that bind a site separate from the nucleotide binding site[41] and display anti-proliferative activity in tumour cell lines[42]. In a previous computational study, we used all-atom molecular dynamics (MD) simulations coupled with dynamic residue networks (DRN) and perturbation response scanning (PRS) to determine allosteric hotspots that may be implicated in modulating conformational dynamics in human Hsp90α[37]. Focusing on the open "v-like" conformation, we identified several residues located at the M-domain:CTD interface that form a central hub in the protein interaction network, and provided evidence that when externally perturbed by random force displacements these residues are capable of selecting global conformational displacements towards the ATPase active closed
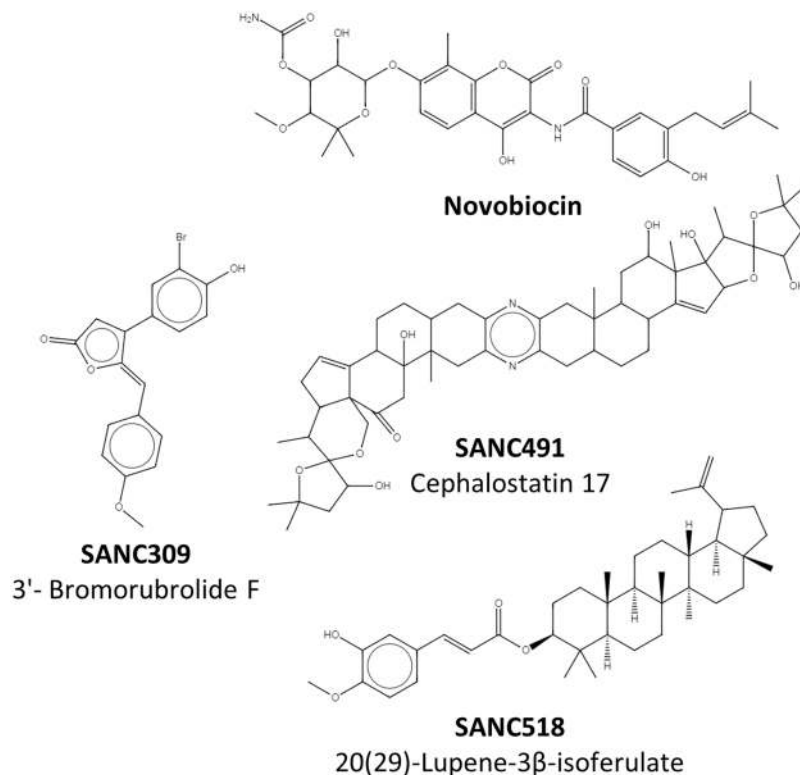
Figure 2. 2D representation of Novobiocin and the identified natural compounds that bind the C-terminal.

conformation. The literature regarding the identification and experimental validation of these allosteric residues was reviewed in detail in our previous study[37].

In the present study we propose that force perturbations at the aforementioned sites, may occur naturally through binding forces engendered by protein-protein and protein-ligand interactions; and that these allosteric hotspots may present a suitable allosteric drug target candidates. We probe two putative ligand binding sites that were identified by combined PRS and DRN analysis[37], and supported by FTMap[43] screening at the CTD as potential allosteric drug targeting sites (Fig. 1 Site-1 & Site-2). We found three South African Natural compounds (SANC) that dock to these sites: Cephalostatin 17 (SANC491) and 20(29)-Lupene-3β-isoferulate (SANC518) preferentially bound at Site-1 together with the known CTD inhibitor Novobiocin which was included for comparative control purposes; and 3′-Bromorubrolide F (SANC309) docked at Site-2. We utilized several all-atom MD simulation based analysis techniques, to assess the allosteric potential of each site to modulate protein conformational dynamics in response to ligand interactions.

Results and Discussion

Identification of novel natural compounds that bind the CTD.
The availability of putative small molecule binding sites at the CTD of the "v-like" open conformation of human Hsp90α[37] was assessed using FTMap[43] with an unbiased screen positioned over the M-domain:CTD region, resulting in two candidate binding sites that overlapped with allosteric control elements identified in our previous study via combined PRS and DRN analysis[37] (Figure S1, Supplementary Material). Site-1 is located at the CTD interface in a groove positioned at the four-helix bundle (helix$_{21,22}$) (Fig. 1-yellow). Site-2 is adjacent to Site-1 and represents a sub-pocket positioned at the M-domain:CTD interface of each protomer (Figs S1, 2A and B, Supplementary Material). Each sub-pocket is formed by residues belonging to helix18, ß-sheet$_{17-18}$, and helix$_{22}$, as well as several residues located at the M-domain:CTD hinge region:

Molecular docking was utilized to separately screen a library of 702 South African natural compounds against Site-1 and Site-2. The known CTD inhibitor Novobiocin was included in both screens as a control. Due to computational constraints, the protein receptor was kept rigid and the ligands allowed flexibility around available rotatable bonds. Each compound was re-docked 100 times to ensure conformational diversity and the results filtered and analysed in terms of conformation cluster size and average estimated free energy of binding. Table 1 provides a summary of the best docked compounds selected for each site.

Of the 702 natural compounds screened, only three putative hits against either CTD binding site were observed, two at Site-1 and only a single hit at Site-2 (Table 1, Fig. 2). In addition to the known Hsp90 inhibitor Novobiocin, Cephalostatin 17 (SANC491) and 20(29)-Lupene-3β-isoferulate (SANC518) reproducibly docked to Site-1. All three ligands recorded similar docking orientations, each compound lining the length of a binding groove positioned over the four-helix bundle with their terminal moieties extending towards the Site-2 sub-pocket in protomer B (Fig. 1 - inset). This observed binding site and orientation is in agreement with

Compound	Name	Site	Cluster size	Average energy (kcal/mol)	Lowest energy (kcal/mol)
Novobiocin	Novobiocin	1	81	−8.75	−8.97
SANC309	3′-Bromorubrolide F	2	79	−9.13	−9.20
SANC491	Cephalostatin 17	1	93	−11.41	−11.44
SANC518	20(29)-Lupene-3β-isoferulate	1	84	−11.03	−11.46

**Table 1.** Summary of the best docked compounds for Site-1 and Site-2.
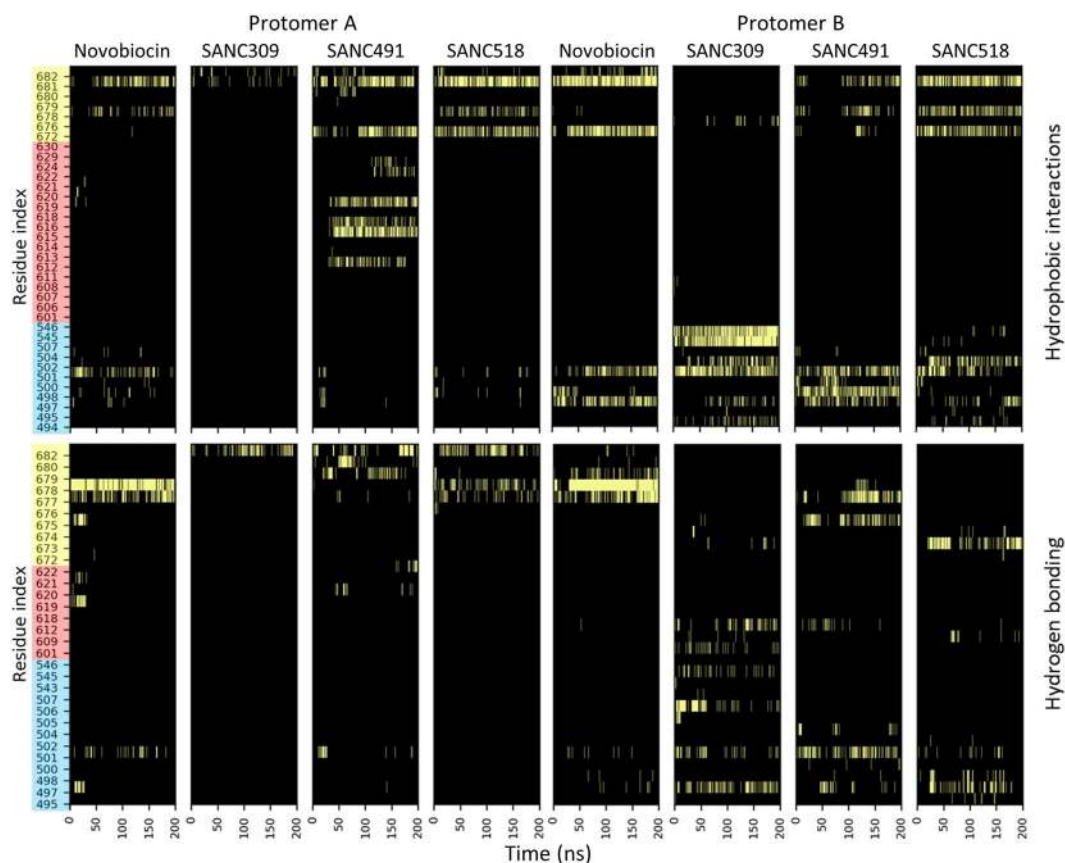
Figure 3. Time evolution of residue contribution to protein-ligand hydrophobic and hydrogen bond interactions. Detected interactions are depicted by light bars. Y-axis residue shading represents the different binding site residues: blue - sub-pocket; red – helix$_{18}$; yellow - four-helix bundle.

previous reports for Novobiocin[44]. Both SANC491 and SANC518 recorded lower binding free energy scores (<-11.0 kcal/mol) compared to Novobiocin (-8.97 kcal/mol) suggesting higher binding affinities. Focusing on Site-2, Novobiocin was unable to access the small sub-pocket in protomer B and only 3′-Bromorubrolide F (SANC309) of the natural compounds was able to dock to the with a high degree of reproducibility and suitable binding free energy scores (-9.13 kal/mol) (Fig. 1 – red). To fully examine the stability of ligand binding and to gain further insights as to the allosteric effect of each ligand, all-atom MD simulations were carried out on the protein-ligand complexes for a total 200 ns. In each case, the ligand conformation with the lowest binding free energy score from the largest cluster was selected as a representative start structure.

Characterization of protein-ligand interactions. Ligand binding stability was assessed by monitoring the residue contribution to protein-ligand interactions over time. Hydrophobic and hydrogen bond interactions were evaluated every nth frame in the trajectory using a 200 ps time interval, and the cumulative data represented as interaction heat maps (Fig. 3). Starting with binding Site-1, Novobiocin, SANC491, and SANC518 primarily interact with residues (L672-Q682) from both protomers that line a distinct binding groove located at four-helix bundle of the CTD inter-protomer dimerization interface (Fig. 3-yellow). Of these residues, L627, S677, L678, and P681 are the most consistent contributors to ligand interactions. In addition to stable hydrophobic (residues L627 and P681) and hydrogen bond interactions (residues S677 and L678), the terminal moieties of the Site-1 ligands also interact with residues (T495-F507) that line the entrance to the sub-pocket binding site in protomer B (Site-2) (Fig. 3-blue). The only notable difference between the Site-1 ligands was the formation of additional
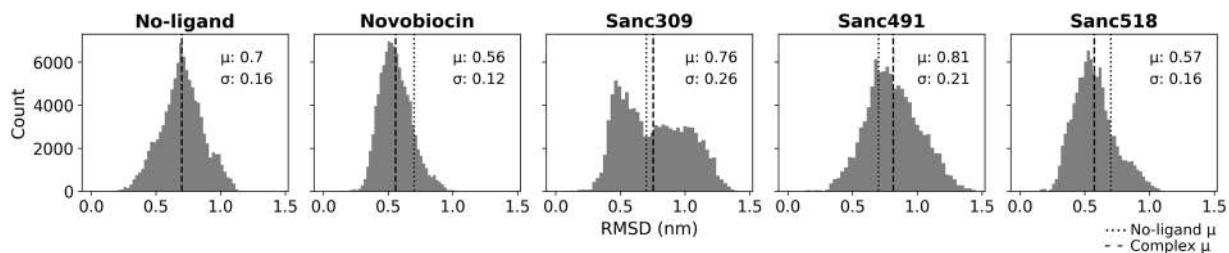
Figure 4. Whole protein backbone RMSD distribution plots. Comparison of the histogram distribution plots for each ligand bound complex with the ligand-free state demonstrates variable conformational flexibility. The shift in the conformational distribution over the 200 ns MD trajectory can be visually assessed by comparing the mean (μ) of each complex (dashed line) with the ligand-free complex (dotted line). σ corresponds to the standard deviation.

interactions between SANC491 and helix$_{18}$ of protomer A (residues R612-T624) (Fig. 3-red). The observed orientation of the binding Site-1 ligands together with their corresponding interacting residues is in close agreement with a previous biochemical and *in silico* study of Bisphenol A based allosteric inhibitors of human Hsp90[42]. Furthermore, interacting residues L672, S674, and P681 are closely positioned to, and overlap with, several CTD allosteric hotspots (residues599-W606, and T669-L678) which have previously been implicated in NTD allosteric signalling and control of conformational dynamics[33].

Looking at binding Site-2, SANC309 appears to interact exclusively with residues belonging to protomer B (residues T495-F507 and S543-K546, Fig. 3-blue) with the exception of hydrogen bond interactions with the four-helix bundle through residue Q682 in protomer A (Fig. 3-red). In protomer B, residues Q501, T545 and K546 form stable hydrophobic interactions with SANC309 while interactions with the remaining sub-pocket residues appear to be more transient (Fig. 3 – blue). The protein-ligand interaction landscape observed for SANC309 is to the best of our knowledge novel to the current study and notably overlaps with several allosteric hotspot residues (T495, E497, T545, and K546) that have been previously implicated in allosteric modulation of conformational displacements in favour of the closed conformation when externally perturbed[37].

Overall, MD simulations revealed stable protein-ligand complexes over 200 ns, and the interaction profiles for both Site-1 and Site-2 overlap with known allosteric sites opening the possibility for external modulation of Hsp90α conformational dynamics through ligand binding interactions. We investigate this possibility by monitoring the effect each ligand has on the global and internal dynamics of the protein compared to a ligand-free system and assess the respective allosteric potential of Site-1 and Site-2.

CTD ligands modulate protein flexibility. Backbone root-mean-square-deviation (RMSD) analysis serves as a measure for monitoring conformational variation over an MD trajectory. Whole protein RMSD analysis indicates that in the ligand-free state, Hsp90 experiences variable backbone flexibility recording RMSD values that are normally distributed around a mean of 0.70 nm with a standard deviation of 0.16 nm (Fig. 4, and Figure S2-black, Supplementary Material), an observation that suggests inherent plasticity of Hsp90 in the open conformation[45]. The addition of either Novobiocin or SANC518 appears to improve the stability of the dimer compared to the ligand-free system, recording normally distributed RMSD values centred on reduced means of 0.56 and 0.57 nm respectively. Similar standard deviations of 0.12 and 0.16 nm for these complexes confirms a degree of variable flexibility for the open conformation. In contrast, the presence of either SANC309 or SANC491 appears to enhance conformational flexibility, each complex recording increased mean RMSD values of 0.76 and 0.81 nm, coupled with increased standard deviations of 0.26 and 0.21 nm respectively (Fig. 4 and Figure S2-black, Supplementary Material). Furthermore, SANC309 records a bimodal RMSD distribution compared to the normal distributions of the four remaining complexes, providing evidence of variable conformational sampling. The statistical significance of these findings is reported in the Supplementary Material Table S1. In summary, the data presented here tentatively suggests an increased flexibility for SANC309, and SANC491 and a maintained if not reduced flexibility for Novobiocin and SANC518.

This observed plasticity of the Hsp90 dimer could be explained by independent conformational restructuring of either protomer. To assess thus we examined the RMSD profiles for each protomer in isolation (Figs S2 and S3 Supplementary Material). With the exception of SANC518, the protomer RMSD distribution for each complex is normally distributed, recoding similar but not identical average and standard deviation values (μ: ±0.32 nm, σ: ±0.06) compared to the ligand-free system, suggesting the protomers are able to move independently of each other within the bounds of CTD dimerization. In contrast, the presence of SANC518 appears to enhance individual promoter flexibility, shifting the mean to 0.43 nm in Protomer A and 0.51 nm in Protomer B, while increasing the standard deviations of each to 0.10 and 0.18 nm respectively (Fig. S3 and Table S1 Supplementary Material). The marked difference in protomer B of this complex is coupled to a bimodal distribution with peaks centred at ~0.38 and ~0.70 nm respectively, suggesting conformational sampling of two dominant structures with the former likely corresponding to that of the ligand-free system.

It is likely that the observed backbone flexibility/mobility of the whole protein dimer may result from conformational repositioning of the semi-rigid protomers relative to one another in either a linear opening/closing manner, or through perpendicular rotation about the C-terminal dimerization axis. The distribution of the inter-protomer distance, defined as the measured distance between the center of mass for each NTD, informs on
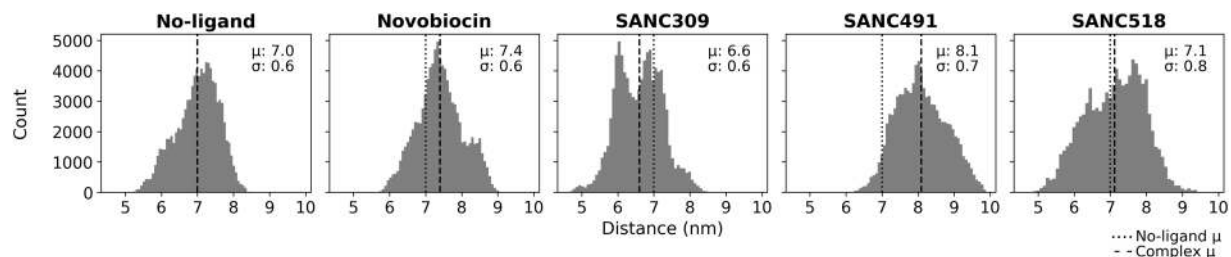
Figure 5. Distribution of inter-protomer distance. Inter-protomer distance is defined as the distance between the center of mass of each NTD. The conformational population shift in each complex is demonstrated by comparing the mean (μ) of each complex (dashed line) with the ligand-free complex (dotted line). σ corresponds to the standard deviation.
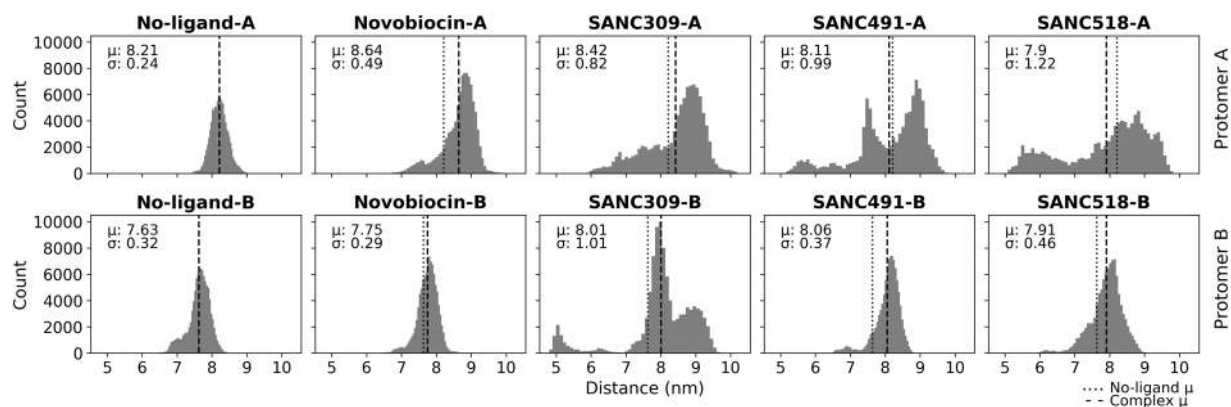

Figure 6. Flexing around the central axis of each protomer. Protomer flexing is measured as the distance between the center of mass of each NTD and the CTD interface. The conformational population shift in each complex is demonstrated by comparing the mean (μ) of each complex (dashed line) with the ligand-free complex (dotted line). σ corresponds to the standard deviation.

the propensity for Hsp90 to populate conformations distinct from the starting structure over the course of the 200 ns MD trajectory (Fig. 5). In the absence of CTD ligands, this distance distribution is positioned around the initial NTD-NTD distance of 7.6 nm, recording a mean distance of 7.0 nm (Fig. 5, and Table S2, Supplementary Material) indicating limited conformational variation under these conditions. For the ligand bound systems, Novobiocin and SANC518 appear to populate conformations analogous to the ligand-free system (Fig. 5) with mean values of 7.4 and 7.1 nm respectively, while addition of SANC309 and SANC491 cause a shift in the distribution away from the ligand-free system. SANC491 populates conformations with an increased average inter-protomer distance of 8.1 nm (Fig. 5–green), suggesting a conformational preference for a more open "v-like" structure. SANC309 on the other hand, shifts the distribution to the left (μ 6.6 nm) and the formation of a bimodal distribution suggests population of two dominant conformations: one analogous to the ligand-free structure (±0.70 nm) and the other in favour of a more closed conformation (±6.0 nm).

Inasmuch as the NTD-NTD distance provides information on the linear opening/closing of the homodimer, the distribution of the NTD-CTD distance for each protomer informs on protomer flexibility around the central axis as seen in hinge bending motions. Shorter NTD-CTD distances could be indicative of protomer bending, while increased NTD-CTD distances could suggest protomer straightening/extension. The NTD-CTD distance is defined as the distance between the centers of mass of the NTD and CTD four-helix bundle (Fig. 1-yellow), and the time evolution of these measurements is represented as a distribution for each protomer separately (Fig. 6).

In the absence of bound ligand, both protomers form distinct distribution peaks around the mean NTD-CTD distance (μ 8.21 nm: protomer A and μ 7.63 nm Protomer B) which correspond to the initial NTD-CTD measurements of 8.0 nm and 7.5 nm, suggesting minimal hinge bending motions and confirming the rigid nature of the protomers under these conditions. Addition of Novobiocin appears to have a differential effect on the observed distributions, which is skewed to the right (±9.0 nm in protomer A and ±7.75 nm in protomer B) when compared to the ligand-free system. This positive shift in the inter-domain distance may be indicative of ligand induced sampling of more extended protomer conformations. Indeed, similar albeit exaggerated observations can be made for the other Site-1 ligands, in which both SANC491 and SANC518 induce skewed bimodal distance distributions in protomer A with dominant peaks centred ±9.0 nm, while protomer B experiences tighter skewed distributions at ±8.1 nm. The formation of alternate peaks in protomer A indicates a wider range of sampling under these conditions providing evidence of ligand induced flexibility around the central axis. Binding of SNAC309 at Site-2 also appears to have a tuneable effect on the flexibility of the protomers when compared to the ligand-free system, particularly protomer B, which forms three distinct distribution peaks (5.0, 8.0, and
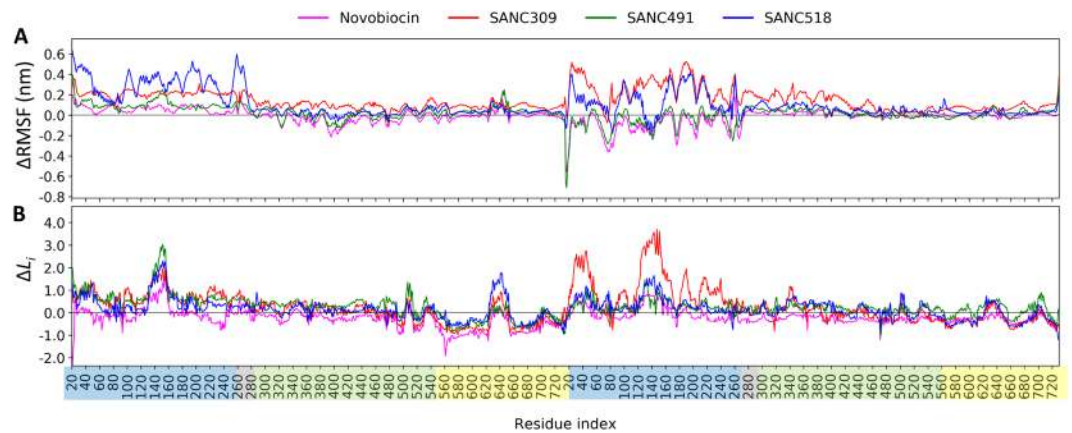
Figure 7. Effect of CTD bound ligands on internal dynamics: (**A**) ΔRMSF, and (**B**) ΔL_i plots for each ligand bound complex, calculated as the average difference between the protein-ligand and ligand-free complexes for each residue. Residue indices are coloured by domain: NTD – blue, M-domain – green, CTD – yellow.

9.0 nm), while protomer A forms a skewed distribution peak at ±9.0 nm. Taken together, this data suggests that ligand interactions at the CTD may allosterically induce protomer flexibility allowing the dimer to sample alternate structural conformations. Comparative statistics between the ligand un/bound complexes are reported as Supplementary Material in Table S3.

Next, we focus on the relative effect of bound ligand on the flexibility of localized regions of the protein. The root-mean-square-fluctuation (RMSF) is a measure of the average positional displacement of each residue over time and indicates flexibility/mobility of residue sites. Here, we monitor the relative change in residue fluctuation (ΔRMSF) between the ligand-free and ligand bound complexes to determine the extent to which bound ligands influence intra-protomer flexibility. Inspection of the ΔRMSF profiles (Fig. 7A) reveals differential ligand specific modulation of domain flexibility, whereby bound ligands appear to modulate the RMSF of entire domains rather than individual residue sites. Novobiocin and SANC491 differentially modulate the RMSF of the NTDs of each protomer, causing increased fluctuations in protomer A and a decrease in protomer B (Fig. 7A – blue shading). Meanwhile, SANC309 and SANC518 appear to increase residue fluctuations at the NTDs of both protomers with large positive ΔRMSF values. In addition, SANC309 also experiences increased fluctuations in the M-domains of both protomers compared to the Site-1 ligands (Fig. 7A – green shading).

The collective results presented in this section demonstrate evidence of selective ligand modulation of both whole protein mobility and domain specific flexibility depending on the identity of the bound ligand. Conformational flexibility and rigid body mobility form the basis of enzymatic catalysis and allosteric modulation[46] and in the case of Hsp90, conformational plasticity is crucial for molecular functionality[5,45]. The increased flexibility experienced by SANC309, especially in the M-domain where co-chaperones bind, may enable the protein to overcome energetic limitations allowing it to explore a larger conformational space, thus aiding its search for the closed catalytically active state. Conversely, reduced protomer flexibility in SANC491 may lead to protomer elongation and a more rigid structure that favours the open conformation.

Ligand modulation of residue interaction networks. Dynamic residue networks (DRNs)[47] were utilized to analyse the effect of ligand binding on residue connectivity over time. DRNs were constructed for each MD trajectory by treating C_β atoms (C_α for glycine) as nodes in the network and connections between nodes established based on a distance cut-off of 6.7 Å (see Methods for details), and the resultant DRNs analysed in terms of average long-range residue reachability (L) and average betweenness centrality (BC). In graph theory, the reachability of a residue is defined as the number of connections required to reach residue i from j using the shortest possible path. The average reachability of a residue (L_i) is thus defined as the average number of steps required to reach residue i from any other residue in the network. The metric BC is related to L in that it is a measure of how often on average a residue is utilized in shortest path navigation and has been previously shown to be an effective measure for the identification of functional residues implicated in intra-protein communication[37,48], protein-ligand[49] and protein-protein binding sites[50,51], as well as non-synonymous SNP analysis[52,53]. To evaluate the relative effect of CTD bound ligand on the connectivity of the protein we compare each protein-ligand complex to the ligand-free ATP-only complex by monitoring the change in L_i (ΔL_i) and BC (ΔBC) over the course of the MD trajectory.

Starting with ΔL_i (Fig. 7B), it is evident that ligand binding leads to an increase in L_i, particularly at the NTDs of both protomers (~1 unit), with the largest increase observed for residues belonging to the ATP-lid (residues 130–160). This observation is notably accentuated in protomer B of the SANC309 complex, with ΔL_i values > 3.0. It has been previously shown that ΔL_i can be influenced in one of two ways: (i) a significant spatial alteration in the local neighbourhood surrounding i can affect the contribution of first and second neighbours to the average number of steps taken[37,54]; (ii) the local network surrounding i remains intact but large conformational changes at distant locations in the protein affect the average path length. We have previously shown that there exists a proportional relationship between RMSF and L_i, when residue fluctuations/displacements exceed the distance threshold used to construct the DRN[37]. This observation is evident in the present study, comparison of RMSF
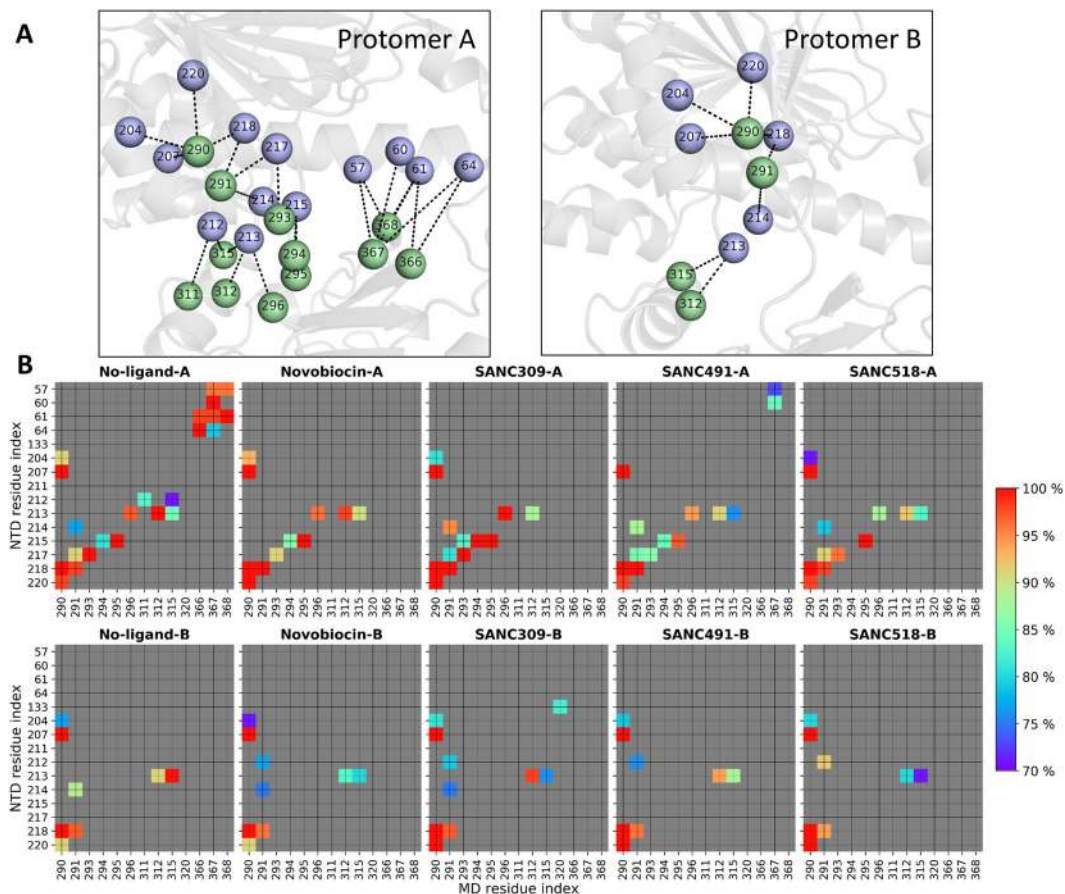
Figure 8. Inter-domain contacts between the NTD and M-domain of each protomer. (**A**) Illustration of the inter-domain contacts between the NTD (violet spheres) and M-domain (green spheres) for the ligand-free ATP-only complex. (**B**) Time progression of key contacts between the NTD and M-domain over the course of each respective MD trajectory. Stable contacts are colored (blue – red) by percent detected over the trajectory and grey if detected <70% of the total simulation time.

and L_i yields strong Pearson's correlation coefficients (>0.85) for all five complexes (Fig. S4A, Supplementary Material). Interestingly, this relationship does not hold true when considering ΔRMSF and ΔL_i, which record significantly poorer Pearson's correlations (<0.7) (Fig. S4B, Supplementary Material). This observation suggests that the positive ΔL_i experienced in the ligand bound systems is not directly linked to increased residue fluctuations as described in scenario (i), and thus we consider scenario (ii) for an explanation for the increase in L_i.

The variable flexing/bending of both protomers in response to bound ligand (Fig. 6), may result in sufficient conformational deformations so as to disrupt established key interdomain network contacts and concomitantly increase the average path length for entire regions of the protein. To investigate this possibility we consider the time progression of all NTD:M-domain contacts using a 200 ps time interval and define interacting residues through a spatial distance cut-off (≤6.7 Å) (Fig. 8). The resulting data set is filtered to interactions that maintain contact for at least 70% of the MD trajectory, and we classify these as stable interactions and color them according to their respective percentage contact duration over time (Fig. 8B). Looking at the ligand-free system, it is evident that there are more stable inter-domain contacts between the NTD and M-domain in protomer A compared to protomer B, however two key inter-domain contact regions are present in both protomers (Fig. 8A): residues K204, V207, I214, I218, and L220 in the NTD form stable contacts with L290 and N291. In addition to these contacts, residues D57, R60, Y61 and L64 in the NTD of protomer A maintain stable inter-domain interactions with R366-V368. For the ligand bound systems, it is evident that the inter-domain interactions through reside L290 are maintained in both protomers, and those involving N291 are weakened in protomer B. Notably, the stable inter-domain interactions between the R366-V368 triad in protomer A and the NTD are lost in all ligand bound complexes with the exception of SANC491 which retains partial interactions between R367 and the NTD residues D57 and R60. Taken together this suggests that breakdown of stable inter-domain interactions, particularly those involving the R366-V368 triad in protomer A and N291 in protomer B, may lead to NTD:M-domain decoupling, which in turn may result in the redirection of inter-domain network communication to pass through the long NTD:M-domain linker and ultimately increase L_i. Indeed, NTD decoupling may be evidenced by the enhanced NTD residue fluctuations observed at the NTD of protomer A (Fig. 7A).

While the NTD and M-domain record elevated ΔL_i, the CTDs experience a ~1 unit decrease in L_i in both protomers. This observation is particularly apparent for residues 550–620 belonging to helix$_{17}$ and helix$_{18}$, as well
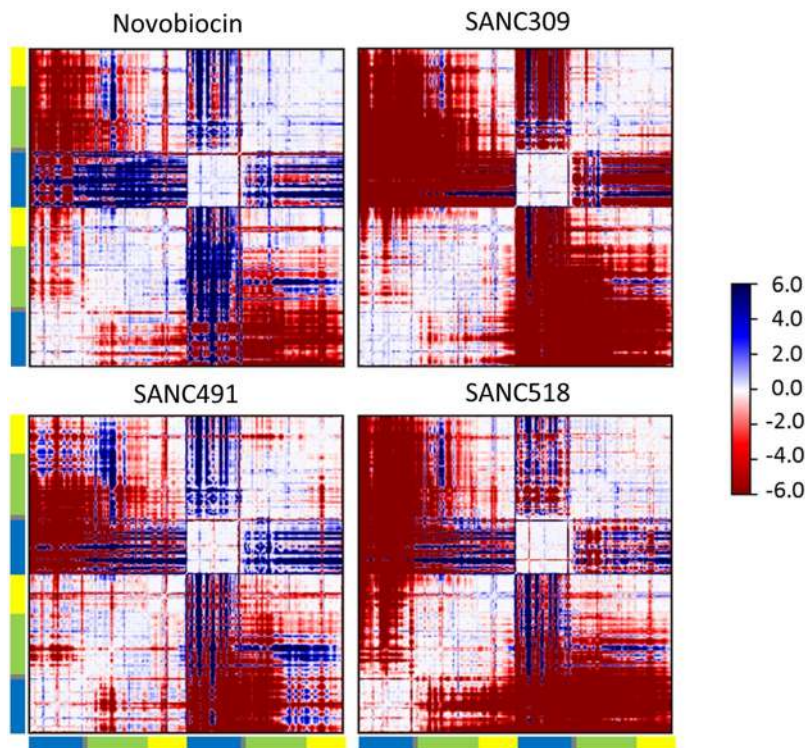
Figure 9. Difference in communication propensity. Matrices calculated as the difference in communication efficiency between the ligand-free complex and the ligand bound complex. Positive values (blue) denote higher communication propensity for the ligand bound complex compared to the ligand-free complex while negative values (red) indicate lower communication propensity for the ligand bound complexes.

as the terminal four-helix bundle (residues 660–720). Interestingly, the former have been previously implicated as sensors to physical perturbations elsewhere in the protein[37] and implicated in signal propagation[33]. Here, scenario (i) provides a suitable explanation in that ligand binding at the CTD stabilizes the domain reducing residue fluctuations thus allowing for stable residue-residue interactions.

As with L_i, ΔBC is also averaged over time and comparison of the different ligand bound complexes reveals a high degree of overlap between the ligand binding sites and regions of the protein that experience large shifts in BC (Fig. S5, Supplementary Material), demonstrating the metric's sensitivity to putative ligand binding sites. Interestingly, all four ligand complexes experience increased BC at the four-helix bundle in protomer A, but a decrease in protomer B. Furthermore, it is evident that the ligand interactions between SANC491 and helix$_{18}$ (red) in protomer A has a direct impact on BC at this location possibly due to ligand stabilisation of this inherently flexible region as demonstrated by the differential ΔL_i values (Fig. 7). Indeed, this finding is in agreement with our previous study in which we report an inverse relationship between BC and L_i[37].

Effect of ligand binding on communication propensity. The methodology for calculating the communication/coordination propensity (CP) residue pair was first introduced for elastic network models by Chennubhotla and Bahar[55] and subsequently extended by Morra *et al.*[33] for MD trajectories. In context of the latter, CP for any two residues describes signal transduction events as a function of the mean-square distance fluctuation between the C_α–C_α atoms over the trajectory. Inter-residue fluctuations that occur at lower intensities are expected to communicate more efficiently compared to that of high intensity C_α–C_α fluctuations[33,55]. It is important to note that CP denotes communication time and thus smaller values represent more efficient commination between any two C_α–C_α atoms. To obtain a compact representation of how bound ligands affect CP, we define the difference matrix between all residue pair CPs determined for the ligand-free and ligand bound systems (Fig. 9). Thus, positive ΔCP (blue) indicates more efficient (faster) communication with the addition of ligand, while negative ΔCP (red) indicates slower communication (slower) in the presence of bound ligand.

Comparing the ΔCP matrices in Fig. 9, we note a similar block character for regions of the protein that display either efficient (+CP) or slower (−CP) coordination propensities. The intra-protomer coordination patterns provide a visual means by which to compare the inter-domain communication propensities. For the Site-1 complexes, the NTD in protomer A appears to be decoupled from the M-domain and CTD as indicated by the red blocks. This decoupling is particularly apparent for SANC518 and resonates with our findings in the previous inter-domain contact analysis (Fig. 8). In protomer B, the opposite effect is observed as indicated by the large blue blocks in which the NTD experiences more efficient CP between the M-domain and CTD, however this observation is less apparent for SANC518 which experiences some inter-domain decoupling in protomer B. Looking at the inter-protomer coordination patterns for these complexes, it is largely evident that ligand binding causes
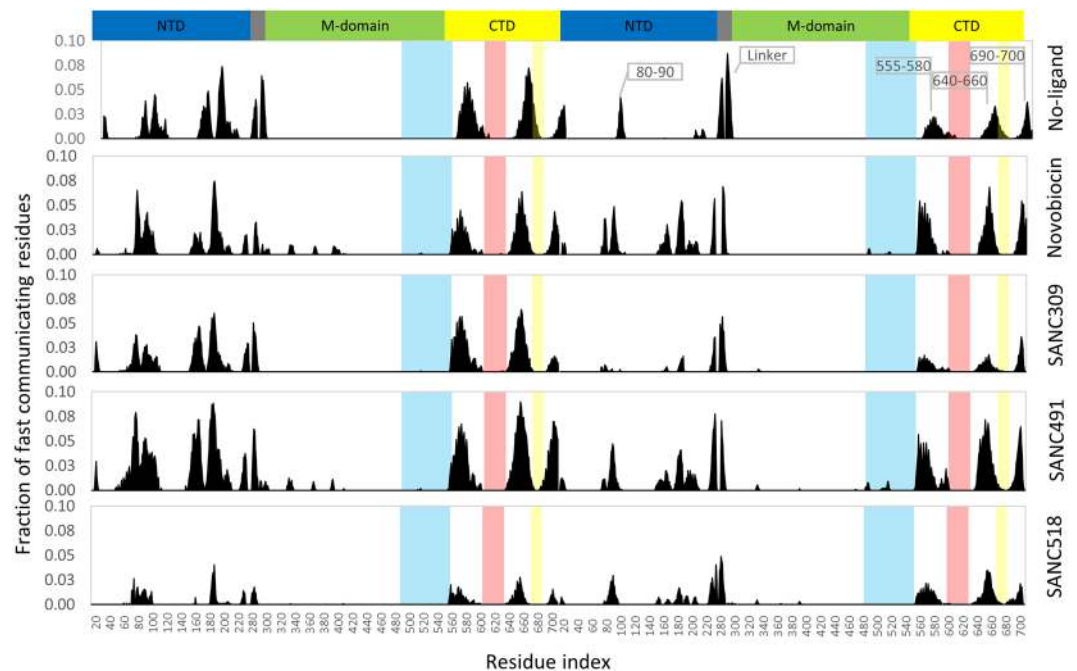
Figure 10. Fraction of fast communicating residues over 80 Å distances demonstrating communication efficiently. Each histogram refers to a single residue and indicates the total fraction of residues that communicate with it. Colored shading represents the ligand binding residues at Site-1 – yellow and red; and Site-2 – blue.

a loss in communication efficiency between protomers. However, in the presence of Novobiocin increased CP is observed between the NTD of protomer B and the M-domain:CTD region of protomer A, while SANC491 experiences a moderate increase in CP between at the NTD of protomer A and the CTD of protomer B. Looking at the coordination patterns SANC309, ligand binding at Site-2 appears to have a drastic effect on CP causing significant intra- and inter-domain decoupling within and between both protomers as evidenced by the large red blocks. It is interesting to note that this observation is particularly apparent in protomer B which provides the sole binding interactions for SANC309. Overall, these observations are consistent with the ΔRMSF results presented in Fig. 7, whereby bound ligands at binding Site-1 appear to increase residue fluctuations at the NTD of protomer A and decrease fluctuations at the NTD of protomer B, while ligand binding at Site-2 increases residue fluctuations throughout both protomers.

We next investigate the relative contribution of individual residues to long range coordination with distant residues elsewhere in the protein. The average CP for neighbouring residues ($i \pm 4$) for the ligand-free complex is 0.85, and we set $CP = 0.85$ as a suitable threshold for discriminating fast communication between residues positioned >80 Å from one another in a similar manner described by Morra and co-workers[33]. By sequentially scanning the protein we record the fraction of residues in the whole protein capable of communicating with a $CP \leq 0.85$ (Fig. 10). Looking at the ligand-free complex it is evident that long range communication is established between the NTDs (residues 80–90, 150–160, 170–185) and CTDs (residues 555–580, 640–655, and 690–700) providing further evidence of NTD-CTD allosteric coupling in agreement with previous reports for yeast Hsp90 in the closed state[33]. For the ligand bound complexes, it is evident that none of the binding site residues directly contribute to long range residue communication (Fig. 10 – shaded areas). Rather, the presence and identity of the bound ligand appears to impact the relative long range coordination of residues in close proximity to the bound ligands. Novobiocin and SANC491 increase the fraction of communicating residues at NTD and CTD of both protomers (residues 80–90, 160–200, 555–580, 640–660 and 690–700). SANC309 appears to significantly reduce the communication efficiency of the CTD, as well as residues 80–90 at the NTD of protomer B, while SANC518 significantly reduces long range coordination in both protomers.

Essential dynamics analysis. Given the tremendous size of Hsp90 (1400 residues), our 200 ns MD trajectories are of an insufficient length to access functional global motions such as the full closing motion expected for the ATP-only complex. Here, essential dynamics (ED) analysis techniques were employed to assess whether ligand binding corresponds to functional global correlated motions. In ED, the MD trajectory is represented as the covariance matrix and the corresponding eigenvectors and eigenvalues used to describe correlated global motion. The former metric represents the correlated displacement of atom groups through essential space, while the latter gives an indication of the magnitude (nm²). The configurational space represented by eigenvectors can be separated into two subspaces[56]; (i) the essential subspace which represents correlated motions comprising very few degrees of freedom, and likely points to biologically relevant global motions; and (ii) the independent subspace which is constrained to local regions and offers little functional importance. Here, we focus on the former subspace (i) which is often accounted for by the first few low-frequency modes with large
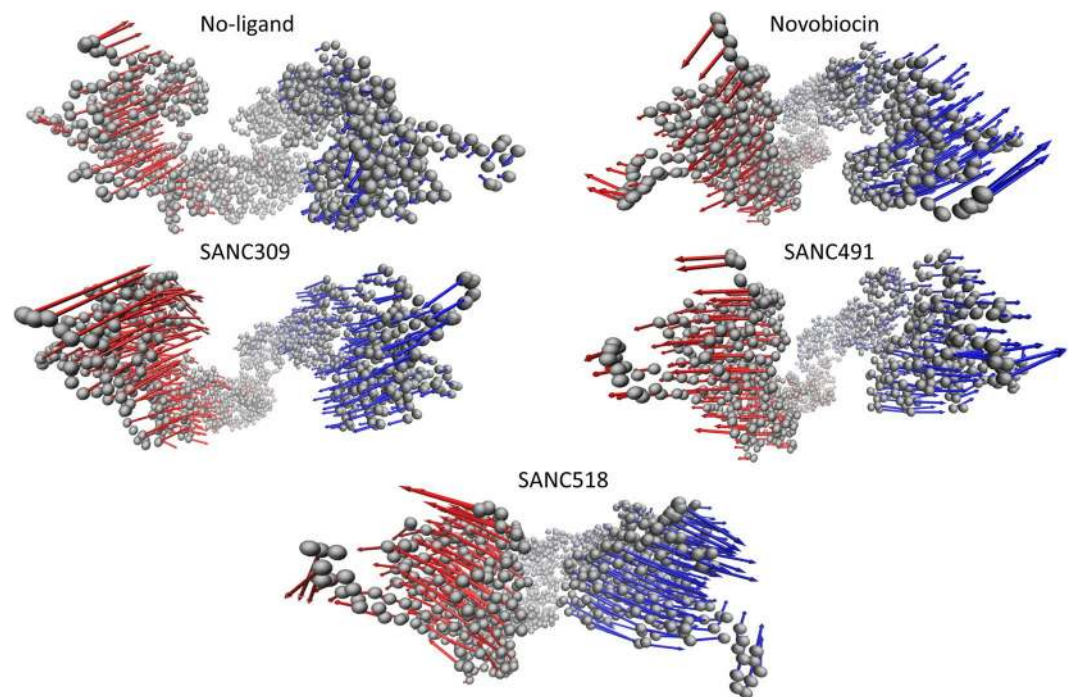
Figure 11. Illustration of the global correlated motion for the first eigenvector of each Hsp90 complex. The protein complex is represented by C_α atoms (grey spheres) and the arrows describe the relative direction and magnitude correlated motion for protomer A (blue) and protomer B (red). Displacement arrows were drawn for every 2nd C_α atom to simplify presentation.

corresponding eigenvalues. Analysis of the cumulative squared overlap for the first 20 modes for each complex (Fig. S6, Supplementary Material), shows that between 60% and 80% of all protein motion is accounted for by the first three eigenvectors. In each case the first two eigenvectors account for more than 50% of all protein fluctuations, suggesting these modes represent functionally relevant protein motions. We illustrate the direction and magnitude of the global correlated motions associated with the first eigenvector for each complex by projecting the corresponding trajectory onto the eigenvector and interpolating over the two most extreme projections, using arrows to describe the relative atomic displacements (Figs 11 and S7, Movies S1–S5, Supplementary Material).

In the ligand-free state, ED analysis reveals linear global correlated motions for each protomer in the opposite direction towards one other, which closely resemble the closing transition that is expected under these conditions, in which ATP binding triggers conformational rearrangements towards the catalytically active closed state. Addition of Novobiocin engenders correlated protomer motions in opposite directions in what can be described as an opening like motion. This motion may provide early evidence of protomer uncoupling as Novobiocin is known to disrupt protomer dimerization at the CTD[32]. Interestingly, SANC491 and SANC518 demonstrate similar linear motions to Novobiocin, providing further evidence that these compounds may invoke a similar inhibitory allosteric mechanism. SANC309 on the other hand appears to enhance global correlated protomer motions towards each other in a similar manner to the ligand-free ATP-only complex. This observation coupled with an increased magnitude compared to the ligand-free complex, may suggest SANC309 to be an allosteric activator of human Hsp90α, promoting allosteric closure of the dimer. Indeed, this observation agrees with earlier results in which SANC309 populated conformations with smaller inter-protomer distances (Fig. 5). Furthermore, we note similar results for the second eigenvectors which describe rotational twisting of the protomers along their central axis in either a clockwise closing or an anti-clockwise opening motions much like the twisting and untwisting of a double stranded helix.

To confirm these findings, we assessed the change in sampling of the conformational space of the protein by comparing the aforementioned correlated motions with known structural reference data. In a previous study we assessed the allosteric effect bound nucleotide has on the conformational dynamics of human Hsp90α. In this study, full-length homology models of Hsp90α in the fully-closed, partially-closed, open, and fully-open states were submitted to long-range (200 ns) all atom MD simulations. Here we utilized 20 ns equilibrated portions of these ATP-bound trajectories to construct a concatenated trajectory totalling 80 ns of Hsp90α in four known conformational states. PCA analysis was conducted on this trajectory to attain the covariance matrix. Analysis of cumulative squared overlap for first 20 modes of this covariance matrix indicates that nearly all global motions are accounted for by the first three eigenvectors (PC1: 77%, PC2: 95%, and PC3: 97%) (Fig. S6-grey, Supplementary Material), confirming these motions to be biologically relevant. Projection of the MD trajectories of the ligand un/bound complexes together with the four known state trajectories onto the first three principle components (PC1: linear opening, PC2: twist closing PC3: linear closing) in a pairwise manner allows for the determination of a change in conformational sampling with respect to known reference states.
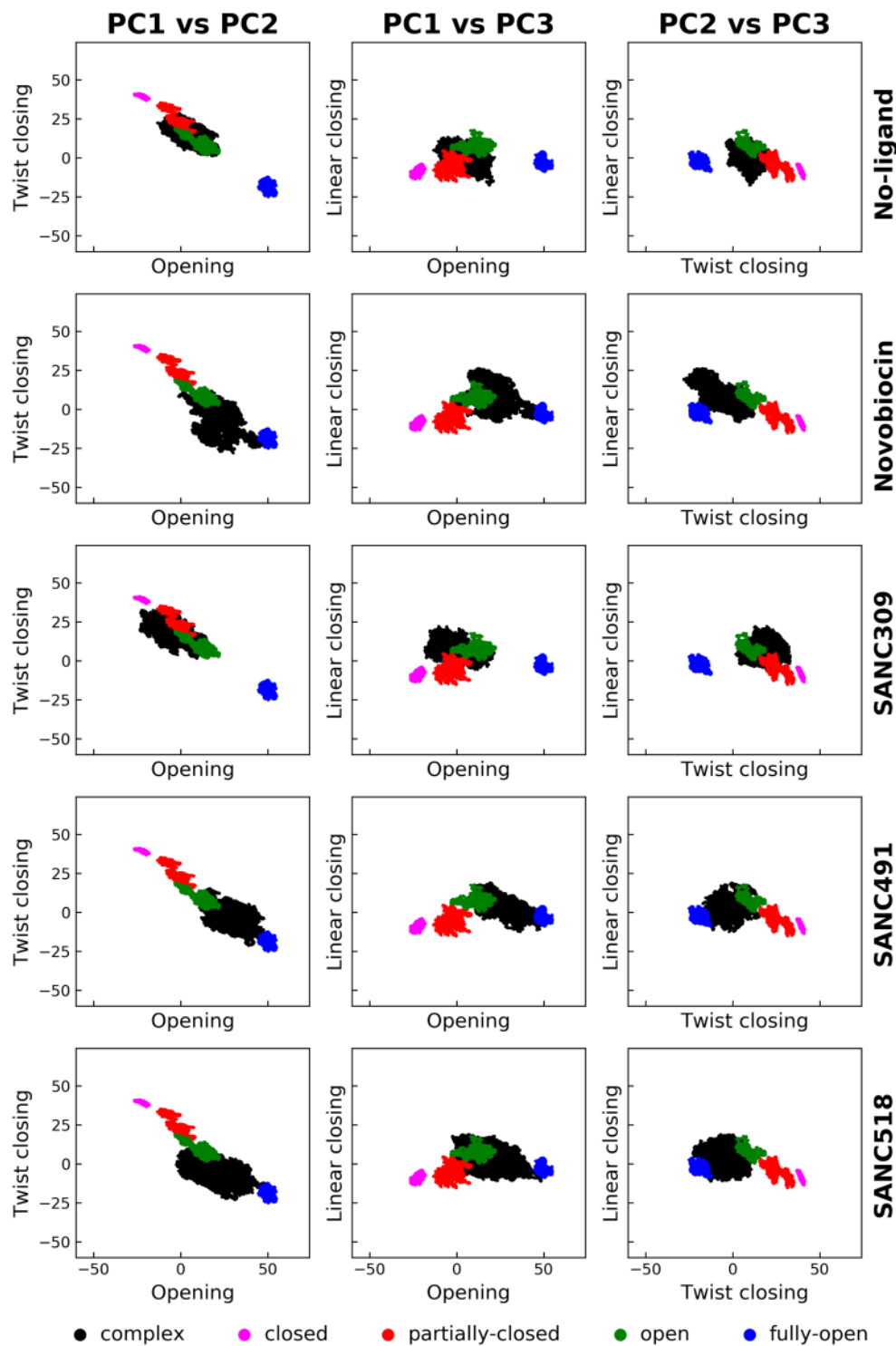
Figure 12. Comparative pairwise score plots demonstrating ligand induced conformational sampling. Each row illustrates the conformational sampling of each ligand un/bound complex (black) in relation to the four known conformational states (see legend). Each column describes the projection of each complex onto the top three eigenvectors (PC1: linear opening, PC2: twist closing, PC3: linear closing).

Comparison of the resultant pairwise scatter plots illustrates the resultant conformational sampling of the ligand un/bound complexes with respect to the known structural states (Fig. 12). We note that the initial starting conformation (Fig. 1) for the ligand un/bound complexes in this study was obtained from the "open state" MD trajectory defined in green in Fig. 12. In the ligand-free state, Hsp90 appears to populate conformations in favour of the closed and partially-closed conformations, confirming the dominant global correlated motions described in the previous section. In contrast, the addition of Site-1 ligands engenders conformational overlap

with the open and fully-open conformational states, confirming global correlated motions in favour of a more open v-like conformation compared to the initial open structure. Finally, analysis of the SANC309 complex reveals conformational sampling that is analogous to the ligand-free system demonstrating significant overlap with partially-closed conformational state, supporting the observed global correlated motions in favour of the closing transition towards the catalytically active conformation.

Conclusions

Overall, our findings establish evidence of ligand specific modulation of the conformational dynamics of human Hsp90α in the open "v-like" conformation. We show how natural forces associated with ligand interactions at two putative druggable CTD binding sites direct the conformational sampling of the dimer by fine tuning the internal dynamics of the protein. Of the binding Site-1 ligands, Novobiocin and SANC491 enhanced conformational rigidity through protomer flexing and reduced NTD residue fluctuations resulting in more efficient NTD-CTD allosteric communication. SANC518, on the other hand, behaved differently enhancing protomer flexibility leading to reduced NTD-CTD allosteric communication in protomer A. Given, the overlap in binding site of the Site-1 ligands with known inhibitors such as Novobiocin and several Bisphenol A based inhibitors, coupled with evidence of correlated opening motions, we propose that small molecules targeting the CTD dimerization site located at the four-helix bundle of the open conformation may externally modulate the conformational dynamics in favour of a more open conformation and thus act as allosteric inhibitors of Hsp90α. These compounds may prevent conformational cycling to the closed catalytically active state by either disrupting CTD dimerization as is the case for Novobiocin[32], or by allosterically enhancing the energetic barrier that must be overcome in order to access the ATPase active state[23]. Finally, we note that Cephalostatin 17 (SANC491) is known to be a potent anti-cancer agent[57,58] although its mechanism of action remains unclear. In contrast, SANC309 at Site-2 greatly enhanced protein flexibility and decreased internal coordination, characteristics that may enable the protein to overcome the energetic requirements necessary to access the closed state and therefore act as an allosteric activator. Indeed, the correlated global atomic motions observed for this complex are in agreement with our previous PRS study, in which external force perturbations at Site-2 resulted in global protein displacements towards the closed catalytic conformation[37], providing proof of concept for this approach through simulated ligand interactions. In summary, our findings provide novel insights regarding the selective external modulation of Hsp90α conformational dynamics, and shed valuable insight on allosteric drug development for Hsp90 as well as the protein's complex allosteric mechanism of action. Additionally, this is the first study that applies combined information from PRS coupled with DNR to identify potential allosteric sites for inhibitor design. This proposed approach should be applicable in identifying allosteric drug targeting sites in other proteins.

Methodology

Molecular docking. The 3-dimensional coordinate data for human Hsp90α in an open "v-like" conformation was obtained from previous MD simulations of a homology model of human Hsp90 in the presence of various nucleotide conditions[37]. Clustering analysis was carried out on a 400 ns all-atom MD trajectory of Hsp90α in complex with ATP according to the methodology described by Daura and co-workers[59], and the frame with the smallest RMSD to the average of the largest cluster selected the representative protein receptor. Natural compounds indigenous to Southern Africa were obtained from SANCDB[60], while the known Hsp90 inhibitor Novobiocin was retrieved for the ZINC database. Both the compounds and the protein were prepared for docking using the AutoDockTools software suite[61]. The previously identified allosteric binding sites were mapped onto two separate grids with AutoGrid4[61]. Site-1 was centred over the four-helix bundle located at the CTD interface (residues G668-A685), and Site-2 on the adjacent pocket located on protomer B (residues H491-V508, V542-Q550, P596-N609). All docking calculations were performed with AutoDock4[61] using the Lamarkian Genetic algorithm with a population size of 150, and the number of evaluations and generations set to 10 000 000. A total of 100 docking runs were performed for each compound, and each run evaluated with the semi-empirical scoring function supplied by AutoDock4. Clustering of the docked conformations was carried out using AutoDockTools with a cut-off of 1.5 Å. Docking of a compound was deemed reproducible if the largest cluster exceeded 50% of the total runs, with an average energy score of less than −8.00 kcal/mol. Due to the novelty of the "v-like" open conformation receptor used in this study, and the lack of experimental structures co-crystalized in complex with inhibitors at either binding site, it was not possible to conduct redocking experiments of known inhibitor ligands to either site and establish a base binding energy score threshold. Instead, we found that the known Hsp90 inhibitor Novobiocin bound the CTD with a similar orientation and binding energy (−9.00 kcal/mol) as reported by Goode *et al.*[42] and set the binding energy threshold to be 1 kcal/mol greater (−8.00 kcal/mol) to increase the change of potential hits at either site. The screened compounds were ranked by average energy score, and the lowest scoring conformation of the best candidates selected as the representative conformation for long range all-atom MD simulations.

Molecular dynamics. All MD simulations were performed using GROMACS 5.1.2[62,63] with in the CHARMM 36 force field[64–66], using a orthorhombic periodic box with a clearance space of 1.5 nm. Water molecules were added as solvent, and modeled with the TIP3P water model, and the system neutralized using a 0.15 M NaCl concentration. Prior to production runs, each system was first energy minimized using a conjugate-gradient and energy relaxed up to 50 000 steps of steepest-descent energy minimization, and terminated when the maximum force <1000.0 kJ/mol/nm. Energy minimization was followed by equilibration, first in the *NVT* ensemble at 310 K using Berendsen temperature coupling, and then in the *NPT* ensemble at 1 atm and 310 K until the desired average pressure (1 atm) was maintained and volumetric fluctuations stabilized. All production simulations were run for a minimum of 200 ns, and the backbone root-mean-square deviation (RMSD) of the protein monitored for convergence. Coordinate data for the protein and ligand were saved at 2 ps intervals for further analysis. All

simulations utilized the LINCS algorithm for bond length constraints and the fast particle mesh Ewald method was used for long-range electrostatic forces. The switching function for van der Waals interactions was set to 1.0 nm and the cutoff to 1.2 nm. NH_3^+ and COO^- groups were used to achieve the zwitterionic form of the protein and periodic boundary conditions were applied in all directions and the simulation time step set to 2 fs. All trajectory clustering analyses were carried out with the GROMACS 5 cluster function using the gromos method described by Daura and co-workers[59] using a cutoff of 0.2 nm. The RMSD values for each complex are displayed in as Supplementary Material Fig. S2.

Protein-ligand interactions. The Protein Ligand Interaction Profiler (PLIP)[17] was used to predict protein-ligand interactions present over the course of each MD simulation. Using a sampling step of 0.1 ns, each trajectory was reduced to 2000 frames, and stored in PDB format. PLIP analysis was then performed on each frame, and in house Python scripts used to parse the PLIP interaction data for each time point and extract those residues involved in either hydrophobic or hydrogen bond interactions with the bound ligand. In this manner the time evolution of bond formation could be assessed.

Dynamic residue networks. To analyse inter- and intra-domain communication, the protein is represented as a residue interaction network (RIN), where the C_β atoms of each residue (C_α for glycine) are treated as nodes within the network, and edges between nodes defined within a distance cut off of 6.7 Å[67]. In this manner, the RIN was constructed as a symmetric $N \times N$ matrix, where the ij^{th} element is assigned as 1 if residue i is connected to residue j and a zero if no connection exists.

In this study, MD-TASK[47] was used to construct dynamic residue networks (DRN) for each MD trajectory, in which RINs are constructed for every n^{th} frame of the trajectory using a 200 ps time interval, to build a DRN matrix. By iterating over the DRN, each RIN is analysed in terms of the average of shortest path length (L_{ij}) between residue i and any other residue j, and betweenness centrality (BC) of each residue. The shortest path length between two residues i and j is defined as being the number of nodes that need to be crossed to reach j from i. The average L_{ij} is then calculated as the average number of steps that the node/residue may be reached from all other residues in the RIN:

$$L_i = \frac{1}{N-1}\sum_{j=1}^{N} L_{ij}$$

(1)

Here, we analyse the change in reachability (ΔL_i) of each residue by monitoring how L_i shifts over the course of the MD trajectory:

$$\Delta L_i = \frac{1}{N}\sum_{m=1}^{N} (L_i^0 - L_i^n)$$

(2)

where L_i^0 denotes the average shortest path length for residue i at time zero, and L_i^n the average shortest path length at frame n.

BC is defined as the number of shortest paths running through a node/residue for a given RIN, and provides a measure of usage frequency each node during navigation of the network. Here, BC was calculated using MD-TASK based on Dijkstra's algorithm[68] and the data rescaled in the range of 0.0–1.0 to be comparable between conformers. Finally, the average BC and ΔL_i are calculated over the DRN, as this measure provides an indication of residues that experience permanent changes in ΔL_i and BC as opposed to minor fluctuations over the course of the MD trajectory[47].

Inter-domain contacts between the NTD and M-domain were evaluated using contact_map.py script from MD-TASK with a distance cut-off of 6.7 Å. The script was utilized every 200 ps and the resultant data collated into a single data frame for plotting purposes.

Communication propensity. The pairwise communication propensity (*CP*) describes the efficiency of communication between residues i and j and is based on the notion that signal transduction events in proteins are directly related to the distance fluctuation of the communicating atoms[33,55]. *CP* is thus defined as a function of the inter-residue distance fluctuations, where residues whose $C\alpha$ -$C\alpha$ distance fluctuates with low intensity are thought to communicate more efficiently (faster) compared to residues whose distance fluctuations are large in which the amplitude of the fluctuations results in slower inter-residue communication. *CP* is calculated as the mean-square fluctuation of the inter-residue distance

$$CP = \left\langle \left(d_{ij} - d_{ij,ave} \right)^2 \right\rangle$$

(3)

where d_{ij} is the time dependent distance between the $C\alpha$ atoms of residues i and j respectively, and the brackets denote the time-average over the trajectory. In this study, the resultant *CP* matrix is used to assess the relative impact each compound has on the internal dynamics and overall flexibility of the protein, by calculating the difference matrix between each ligand bound complex to the ligand-free (ATP only) complex.

Principal component analysis. Principal completed analysis (PCA) was used for the analysis of global motions present over the course of several MD trajectories[56]. This technique involves two main steps: 1) the construction the covariance matrix, **C**, based on the positional deviation of each $C\alpha$ atom, and 2) dimensionality reduction of **C** by diagonalization, to obtain the eigenvectors and eigenvalues. Each $3N \times N$ covariance matrix

was calculated based on an ensemble of protein structures obtained from the respective MD simulation and the elements of **C** defined as

$$C = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$$

(4)

where x_i and x_j are atomic coordinates of each Cα atom, and the brackets denote the average. Eigenvectors with the largest eigenvalues are representative of the slowest modes, and generally are associated with large-scale movements in proteins, which are responsible for protein function. All PCA analyses were conducted using the GROMACS 5 software suite. The covar function was used for the construction and diagonalization of **C**, not including bound ligands and excluding the first 10 ns of trajectory data to avoid equilibration artefacts. The anaeig function was used to project the MD trajectories onto the main eigenvectors.

For the conformational sampling experiment, PCA analysis was conducted using a concatenated trajectory comprised of 20 ns equilibrated trajectory segments of human Hsp90α in the fully-closed, partially-closed, open, and fully-open conformations obtained from previous MD simulations[37]. The 20 ns trajectory segments were concatenated using the GROMACS 5 trjcat program, and PCA analysis conducted on the resultant trajectory as described above.

Data Availability

All data is presented either in the manuscript or in the Supplementary Information.

References

1. Taipale, M., Jarosz, D. F. & Lindquist, S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* **11**, 515–528 (2010).
2. Young, J. C., Moarefi, I. & Hartl, F. U. Hsp90: a specialized but essential protein-folding tool. *J. Cell Biol.* **154**, 267–73 (2001).
3. McClellan, A. J. *et al.* Diverse Cellular Functions of the Hsp90 Molecular Chaperone Uncovered Using Systems Approaches. *Cell* **131**, 121–135 (2007).
4. Zhao, R. *et al.* Navigating the Chaperone Network: An Integrative Map of Physical and Genetic Interactions Mediated by the Hsp90 Chaperone. *Cell* **120**, 715–727 (2005).
5. Schopf, F. H., Biebl, M. M. & Buchner, J. The HSP90 chaperone machinery. *Nat. Rev. Mol. Cell Biol.* **18**, 345–360 (2017).
6. Wayne, N. & Bolon, D. N. Dimerization of Hsp90 is required for *in vivo* function: Design and analysis of monomers and dimers. *J. Biol. Chem.* **282**, 35386–35395 (2007).
7. Mayer, M. P. & Le Breton, L. Hsp90: Breaking the Symmetry. *Mol. Cell* **58**, 8–20 (2015).
8. Prodromou, C. *et al.* Identification and Structural Characterization of the ATP/ADP-Binding Site in the Hsp90 Molecular Chaperone. *Cell* **90**, 65–75 (1997).
9. Meyer, P. *et al.* Structural and Functional Analysis of the Middle Segment of Hsp90: Implications for ATP Hydrolysis and Client Protein and Cochaperone Interactions. *Mol. Cell* **11**, 647–658 (2003).
10. Ali, M. M. U. *et al.* Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. *Nature* **440**, 1013–7 (2006).
11. Harris, S., Shiau, A. & Agard, D. The crystal structure of the carboxy- terminal dimerization domain of htpG, the Escherichia coli Hsp90, reveals a potential substrate binding site. *Structure* 1087–1097 (2004).
12. Hainzl, O., Lapina, M. C., Buchner, J. & Richter, K. The charged linker region is an important regulator of Hsp90 function. *J. Biol. Chem.* **284**, 22559–67 (2009).
13. Tsutsumi, S. *et al.* Hsp90 charged-linker truncation reverses the functional consequences of weakened hydrophobic contacts in the N domain. *Nat. Structual Mol. Biol.* **16**, 1141–1147 (2009).
14. Tsutsumi, S. *et al.* Charged linker sequence modulates eukaryotic heat shock protein 90 (Hsp90) chaperone activity. *Proc. Natl. Acad. Sci. USA* **109**, 2937–42 (2012).
15. Jahn, M. *et al.* The charged linker of the molecular chaperone Hsp90 modulates domain contacts and biological function. *Proc. Natl. Acad. Sci. USA* **111**, 17881–6 (2014).
16. Shiau, A. K., Harris, S. F., Southworth, D. R. & Agard, D. A. Structural Analysis of *E. coli* hsp90 reveals dramatic nucleotide-dependent conformational rearrangements. *Cell* **127**, 329–40 (2006).
17. Dollins, D. E., Immormino, R. M. & Gewirth, D. T. Structure of Unliganded GRP94, the Endoplasmic Reticulum Hsp90: basis for nucleotide induced conformational change. *J. Biol. Chem.* **280**, 30438–30447 (2005).
18. Southworth, D. R. & Agard, D. A. Client-loading conformation of the Hsp90 molecular chaperone revealed in the cryo-EM structure of the human Hsp90:Hop complex. *Mol. Cell* **42**, 771–81 (2011).
19. Prodromou, C. The 'active life' of Hsp90 complexes. *Biochim. Biophys. Acta* **1823**, 614–23 (2012).
20. Scheibel, T. *et al.* ATP-binding properties of human Hsp90. *J. Biol. Chem.* **272**, 18608–13 (1997).
21. McLaughlin, S. H., Smith, H. W. & Jackson, S. E. Stimulation of the weak ATPase activity of human hsp90 by a client protein. *J. Mol. Biol.* **315**, 787–98 (2002).
22. Dollins, D. E., Warren, J. J., Immormino, R. M. & Gewirth, D. T. Structures of GRP94-nucleotide complexes reveal mechanistic differences between the hsp90 chaperones. *Mol. Cell* **28**, 41–56 (2007).
23. Simunovic, M. & Voth, G. A. Molecular and Thermodynamic Insights into the Conformational Transitions of Hsp90. *Biophys. J.* **103**, 284–292 (2012).
24. Hellenkamp, B., Wortmann, P., Kandzia, F., Zacharias, M. & Hugel, T. Multidomain structure and correlated dynamics determined by self-consistent FRET networks. *Nat. Methods* **14**, 174–180 (2016).
25. Schulze, A. *et al.* Cooperation of local motions in the Hsp90 molecular chaperone ATPase mechanism. *Nat. Chem. Biol.* **12**, 628–635 (2016).
26. Whitesell, L. & Lindquist, S. L. HSP90 and the chaperoning of cancer. *Nat. Rev. Cancer* **5**, 761–772 (2005).
27. Jhaveri, K., Taldone, T., Modi, S. & Chiosis, G. Advances in the clinical development of heat shock protein 90 (Hsp90) inhibitors in cancers. *Biochim. Biophys. Acta* **1823**, 742–55 (2012).
28. Taldone, T., Ochiana, S. O., Patel, P. D. & Chiosis, G. Selective targeting of the stress chaperome as a therapeutic strategy. *Trends Pharmacol. Sci.* **35**, 592–603 (2014).
29. Garg, G., Khandelwal, A. & Blagg, B. S. J. Anticancer Inhibitors of Hsp90 Function: Beyond the Usual Suspects. *Adv. Cancer Res.* **129**, 51–88 (2016).
30. Marcu, M. G., Schulte, T. W. & Neckers, L. Novobiocin and related coumarins and depletion of heat shock protein 90-dependent signaling proteins. *J. Natl. Cancer Inst.* **92**, 242–8 (2000).
31. Marcu, M. G., Chadli, A., Bouhouche, I., Catelli, M. & Neckers, L. M. The Heat Shock Protein 90 Antagonist Novobiocin Interacts with a Previously Unrecognized ATP-binding Domain in the Carboxyl Terminus of the Chaperone. *J. Biol. Chem.* **275**, 37181–37186 (2000).

32. Allan, R. K., Mok, D., Ward, B. K. & Ratajczak, T. Modulation of chaperone function and cochaperone interaction by novobiocin in the C-terminal domain of Hsp90: evidence that coumarin antibiotics disrupt Hsp90 dimerization. *J. Biol. Chem.* **281**, 7161–71 (2006).

33. Morra, G., Verkhivker, G. G. & Colombo, G. Modeling signal propagation mechanisms and ligand-based conformational dynamics of the Hsp90 molecular chaperone full-length dimer. *PLoS Comput. Biol.* **5**, e1000323 (2009).

34. Morra, G., Potestio, R., Micheletti, C. & Colombo, G. Corresponding functional dynamics across the Hsp90 chaperone family: Insights from a multiscale analysis of MD simulations. *PLoS Comput. Biol.* **8**, e1002433 (2012).

35. Moroni, E., Zhao, H., Blagg, B. S. J. & Colombo, G. Exploiting Conformational Dynamics in Drug Discovery: Design of C-Terminal Inhibitors of Hsp90 with Improved Activities. *J. Chem. Inf. Model.* **54**, 195–208 (2014).

36. Rehn, A. *et al.* Allosteric Regulation Points Control the Conformational Dynamics of the Molecular Chaperone Hsp90. *J. Mol. Biol.* **428**, 4559–4571 (2016).

37. Penkler, D. L., Atilgan, C. & Tastan Bishop, Ö. Allosteric Modulation of Human Hsp90α Conformational Dynamics. *J. Chem. Inf. Model.* **58**, 383–404 (2018).

38. Morra, G. *et al.* Dynamics-Based Discovery of Allosteric Inhibitors: Selection of New Ligands for the C-terminal Domain of Hsp90. *J. Chem. Theory Comput.* **6**, 2978–89 (2010).

39. Sattin, S. *et al.* Activation of Hsp90 Enzymatic Activity and Conformational Dynamics through Rationally Designed Allosteric Ligands. *Chem. - A Eur. J.* **21**, 13598–13608 (2015).

40. Vettoretti, G. *et al.* Molecular Dynamics Simulations Reveal the Mechanisms of Allosteric Activation of Hsp90 by Designed Ligands. *Sci. Rep.* **6**, 23830 (2016).

41. Roy, S. S. & Kapoor, M. In silico identification and computational analysis of the nucleotide binding site in the C-terminal domain of Hsp90. *J. Mol. Graph. Model.* **70**, 253–274 (2016).

42. Goode, K. M. *et al.* Targeting the Hsp90 C-terminal domain to induce allosteric inhibition and selective client downregulation. *Biochim. Biophys. Acta - Gen. Subj.* **1861**, 1992–2006 (2017).

43. Kozakov, D. *et al.* The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protoc.* **10**, 733–755 (2015).

44. Ishida, R. *et al.* Cisplatin differently affects amino terminal and carboxyl terminal domains of HSP90. *FEBS Lett.* **582**, 3879–3883 (2008).

45. Mayer, M. P. Gymnastics of molecular chaperones. *Mol. Cell* **39**, 321–31 (2010).

46. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).

47. Brown, D. K. *et al.* MD-TASK: a software suite for analyzing molecular dynamics trajectories. *Bioinformatics* **17**, 2768–2771 (2017).

48. Liu, Y. & Bahar, I. Toward understanding allosteric signaling mechanisms in the ATPase domain of molecular chaperones. *Pac. Symp. Biocomput.* 269–80 (2010).

49. Liu, R., Hu, J., Thornton, J., Singh, M. & Funkhouser, T. Computational Prediction of Heme-Binding Residues by Exploiting Residue Interaction Network. *PLoS One* **6**, e25560 (2011).

50. del Sol, A., Fujihashi, H. & O'Meara, P. Topology of small-world networks of protein-protein complex structures. *Bioinformatics* **21**, 1311–1315 (2005).

51. Cheng, T. M.-K., Blundell, T. L. & Fernandez-Recio, J. pyDock: Electrostatics and desolvation for effective scoring of rigid-body protein-protein docking. *Proteins Struct. Funct. Bioinforma.* **68**, 503–515 (2007).

52. Brown, D. K., Sheik Amamuddy, O. & Tastan Bishop, Ö. Structure-Based Analysis of Single Nucleotide Variants in the Renin-Angiotensinogen Complex. *Glob. Heart* **12**, 121–132 (2017).

53. Brown, D. K. & Tastan Bishop, Ö. Role of Structural Bioinformatics in Drug Discovery by Computational SNP Analysis. *Glob. Heart* **12**, 151–161 (2017).

54. Atilgan, A. R., Turgut, D. & Atilgan, C. Screened nonbonded interactions in native proteins manipulate optimal paths for robust residue communication. *Biophys. J.* **92**, 3052–62 (2007).

55. Chennubhotla, C., Bahar, I., Maritan, A., McIntosh, N. & Ziser, L. Signal Propagation in Proteins and Relation to Equilibrium Fluctuations. *PLoS Comput. Biol.* **3**, e172 (2007).

56. Amadei, A., Linssen, A. B. M. & Berendsen, H. J. C. Essential dynamics of proteins. *Proteins Struct. Funct. Genet.* **17**, 412–425 (1993).

57. Pettit, G. R., Xu, J.-P., Schmidt, J. M. & Boyd, M. R. Isolation and structure of the exceptional Pterobranchia human cancer inhibitors cephalostatins 16 and 17. *Bioorg. Med. Chem. Lett.* **5**, 2027–2032 (1995).

58. Pettit, G. R., Xu, J.-P., Chapuis, J.-C. & Melody, N. The Cephalostatins. 24. Isolation, Structure, and Cancer Cell Growth Inhibition of Cephalostatin 20. *J. Nat. Prod.* **78**, 1446–1450 (2015).

59. Daura, X. *et al.* Peptide Folding: When Simulation Meets Experiment. *Angew. Chemie Int. Ed.* **38**, 236–240 (1999).

60. Hatherley, R. *et al.* SANCDB: a South African natural compound database. *J. Cheminform.* **7**, 29 (2015).

61. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–91 (2009).

62. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

63. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).

64. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

65. Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25**, 1400–1415 (2004).

66. Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).

67. Atilgan, A. R., Akan, P. & Baysal, C. Small-World Communication of Residues and Significance for Protein Dynamics. *Biophys. J.* **86**, 85–91 (2004).

68. Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Math.* **1**, 269–271 (1959).

Acknowledgements

Author Contributions

D.P. and Ö.T.B. conceived the study. D.P. designed the study details and wrote the scripts. D.P. analysed the results under the supervision of Ö.T.B. D.P. wrote the initial draft manuscript. Ö.T.B. and D.P. revised and contributed towards the final manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-35835-0.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.