Potential conformational heterogeneity of p53 bound to S100B(ββ)

Chester McDowell, Jianlin Chen and Jianhan Chen

Published Version Information

Citation: McDowell, C., Chen, J., & Chen, J. (2013). Potential conformational heterogeneity of p53 bound to S100B(ββ). Journal of Molecular Biology, 425(6), 999-1010.

Digital Object Identifier (DOI): doi:10.1016/j.jmb.2013.01.001

Publisher's Link: http://www.sciencedirect.com/science/article/pii/S0022283613000041

Potential Conformational Heterogeneity of p53 bound to S100B(ββ)

Chester McDowell[1], Jianlin Chen[2] and Jianhan Chen[1]*

[1]Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA

[2]Department of Hematology, The Central Hospital of Taizhou, Taizhou, Zhejiang, 318000, P. R. China

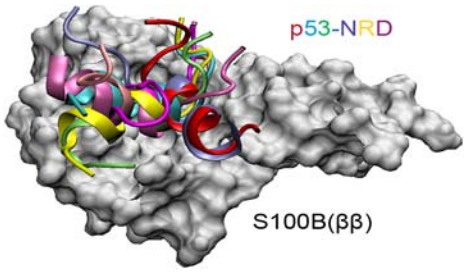
Submitted to *J. Mol. Biol.* as a *Full Length Article*

Revised Version, 2nd

* Corresponding author: Phone: (785) 532-2518; Fax: (785) 532-7278; Email: jianhanc@ksu.edu

Abstract

The negative regulatory domain (NRD) of the p53 tumor suppressor is intrinsically disordered. It contains several post-translational modification (PTM) sites that are important for regulation of p53 activity. Calcium-dependent binding of dimeric S100B(ββ) to p53-NRD blocks access to these PTM sites and disrupts the p53 tetramer to inhibit p53 activation. Previous NMR structural studies have suggested that p53-NRD folds into a stable helix upon binding to S100B(ββ). Intriguingly, despite the well-converged and stably folded nature of the NMR structure ensemble, experimentally resolved intermolecular NOEs are extremely weak; most have 5−6 Å upper bounds, and mainly involve the C-terminal segment of p53-NRD. Such a systematic lack of strong intermolecular NOEs could suggest that the p53/S100B(ββ) interface is more dynamic than currently believed. Indeed, extensive atomistic simulations in explicit solvent (with 1.0 μs total effective sampling) revealed large heterogeneity in the S100B(ββ)-bound conformation of p53-NRD. Helix unwinding at the C-terminus allows key hydrophobic residues (Leu383 and Phe385) to make more extensive intermolecular contacts, whereas the highly helical N-terminus displays substantial flexibility in packing with S100B(ββ). Importantly, the predicted heterogeneous ensemble as a whole is highly consistent with experimental intermolecular NOEs, although many conformational sub-states coexist and individual sub-states satisfy only subsets of the NOE restraints. Furthermore, the simulated ensemble provides similar shielding of key PTM sites to support p53 inhibition. This study not only provides new insights into the structural basis of the p53/S100B(ββ) recognition, but also highlights the importance of recognizing dynamic complexes in structural studies of IDP interactions.



Introduction

Cellular signaling and regulation are frequently mediated by proteins that lack stable tertiary structures under physiological conditions[1-3]. Such intrinsically disordered protein (IDPs) tend to be enriched with charged and polar residues and at the same time lack large hydrophobic residues that are needed for independent folding[4]. As a result, the native states of IDPs correspond to ensembles of heterogeneous and dynamic conformations instead of stably folded structures. Sequence analysis has revealed that intrinsic disorder is highly prevalent in proteomes, with about one-third of eukaryotic proteins predicted to contain disordered segments of >40 residues[5]. The prevalence of intrinsic disorder suggests that conformational heterogeneity (and flexibility) confer important functional advantages, especially in cellular signaling and regulation. Many potential benefits of intrinsic disorder have been discussed, such as larger binding surface areas, inducibility by posttranslational modifications (PTMs), and structural plasticity for binding multiple partners[6-9]. Interestingly, regulatory IDPs frequently fold into stable structures upon specific binding[10]. Such coupled binding and folding allows weaker binding to be achieved without compromising specificity, as the entropic cost of folding offsets the stabilization effects of forming many specific contacts. A reduced binding affinity, in turn, allows higher dissociation rates, which could be an important advantage in signaling and regulation[11].

Coupled binding and folding provides a convenient framework for understanding the structural basis of IDP recognition and regulation, and it has been one of the primary focuses in IDP studies[10;12;13]. Yet, examples have began to emerge where substantial conformational heterogeneity, and sometimes full disorder of the entire binding domain, persists in the bound states[14-16]. This led Tompa and Fuxreiter to suggest that the concept of intrinsic disorder needs to be extended to include protein complexes, arguing that "fuzziness" in the bound state could add "adaptability, versatility and reversibility to the binding of proteins and thereby an ease of regulation in protein-protein interactions"[14]. Such fuzzy complexes could include "static" and/or "dynamic" disorder[14]. A complex with static disorder, such as Tcf4/β-catenin[17], samples a multitude of alternative folded conformations; whereas a complex with "dynamic" disorder contains unstructured segments, either as linkers between folded binding domains ("clamp" model) or as flanking tails of a globular binding domain ("flanking" model). In extreme cases, the entire binding domain could remain disordered in the bound state, and recognition relies completely on transient contacts. Examples of such "random" or "disordered" complexes include

the oligomeric state of the cytoplastimic domains of T-cell receptor ζ–chains[18] and the Sic1/Cdc4 complex[19;20]. Several examples also exist where biological activities appear to be mediated by interactions free of primary sequence constraints[14]. The implication is that these complexes are disordered and do not rely on strict complementarity at the binding interface.

The recognition of the potential prevalence of disorder in protein complexes is an important one. It requires broader consideration of how specific interactions of IDPs might be realized and regulated. For complexes with dynamic disorder[14], establishing the structural basis of recognition becomes more involved, beyond obtaining their high-resolution structures using either X-ray crystallography or nuclear magnetic resonance (NMR); instead, conformational ensembles are required. Unfortunately, resolving heterogeneous protein conformational ensembles is a challenging problem that is often fundamentally limited by underdetermination (that is, the number of experimental restraints is insufficient to uniquely define the accessible conformational space)[21-23]. On the other hand, neglecting potential conformational heterogeneity in the bound state, such as due to lack of awareness of fuzzy complexes, could lead to misleading structural interpretations. It can be particularly problematic for ensemble methods such as NMR and Förster resonance energy transfer (FRET), where a set of independent structural restraints are first determined and then used to calculate multiple structures using restrained molecular dynamics protocols[24]. Neglecting potential disorder during structural calculation will generate overly ordered conformations with structural features that do not necessarily co-exist[23].

The tumor suppressor p53 regulates many genes involved in cell cycle and apoptosis and is arguably one of the most important proteins in cancer[25]; for example, over 50% of human cancers involve some alteration of p53 activities[26]. At the same time, p53 is also one of the most extensively studied IDPs, with multiple intrinsically disordered regulatory domains that mediate p53's interactions with many other proteins[27]. The C-terminal negative regulatory domain (NRD) (residues 367−392), in particular, is one of the few IDPs that have been experimentally shown to be able to fold into distinct structures upon specific binding to different targets, including α-helix, β-strand and two distinct loops[9]. One of the structures involves binding to Ca^{2+}-loaded dimeric S100B(ββ), which blocks access to several phosphorylation and acetylation sites in p53-NRD and thus inhibits p53 oligomerization and transcriptional activity[28]. The level of S100B protein is correlated with malignant melanoma, and is often used as a diagnostic cancer marker[29]. The p53/S100B interaction can also be targeted by small molecule inhibitors for restoring p53 activity in cancer therapeutics[30;31].

A previous NMR structural study[32] suggested that p53-NRD folded into a short helix and formed a structurally stable complex when bound to S100B(ββ) (see Fig. 1). For example, the averaged root mean square deviations (RMSDs) of the NMR structure ensemble (PDB: 1dt7) from the mean structure are only 0.61 Å for all ordered backbone atoms and 1.44 Å when including all ordered heavy atoms[32]. Curiously, most assigned intermolecular nuclear Overhauser effects (NOEs) are extremely weak. Out of a total of 33 unique intermolecular NOEs (see Table 1), 20 were assigned an upper bound of 6 Å and 11 others an upper bound of 5 Å. There are only two medium/weak intermolecular NOEs assigned, both involving p53 Leu383. Closer examination of the NMR ensemble revealed significant uncertainty in side chain packing at the interface, with only a few intermolecular contacts consistently observed within the 40-member ensemble (e.g., see Fig. 2A and Table S1). It also appears that a stably formed helix spanning p53 residues 376–387 would hinder formation of extensive hydrophobic contacts between p53 nonpolar side chains (mainly Leu383 and Phe385) with the narrow but deep hydrophobic binding pocket of S100B(ββ) (see Fig. 1). It was suggested that Arg379 and Lys386 of p53 could form salt bridges with Glu45 and Glu86 of S100B(ββ), respectively[32]; however, exposed salt-bridges are rarely stabilizing[33]. In fact, p53 Lys386 and S100B(ββ) Glu86 are spatially close enough to make salt-bridge contact only ~15% of the time in the NMR ensemble (see Fig. 2A). These observations together suggest that the p53/S100B(ββ) complex structure might not be as stable as previously thought. Interestingly, different levels of conformational instability of p53 have been observed consistently in recent atomistic simulations of the complex[31;34;35] using different force fields including Amber ff99[36], GROMOS 53a6[37] and GBSW implicit solvent[38;39], although force field quality has been mainly invoked to rationalize such observations.

To further examine potential fuzziness and structural basis of the p53-S100B(ββ) interaction, we performed extensive all-atom explicit solvent simulations using two of the best-tested empirical protein force fields, namely, CHARMM22/CMAP[40-42] and Amber ff99SB[43]. Initiated from diverse members of the NMR ensemble, five independent simulations of the complex were carried out, allowing a total of 1.0 μs effective sampling of the p53/S100B(ββ) interface. The resulting conformational ensemble supports significant heterogeneity in p53 conformation when bound to S100B(ββ). Interestingly, such a heterogeneous ensemble appears to be highly consistent with experimentally assigned intermolecular NOEs. Clustering analysis further demonstrated that many conformational sub-states co-existed in the bound state and each satisfied a different sub-set of NOEs. Importantly, the proposed fuzzy complex structure

ensemble appears to be capable of providing similar shielding of key p53 PTM sites and thus could support the inhibition of p53 activation via S100B(ββ) association.

Results

Conformational heterogeneity of the p53/S100B(ββ) interface

The p53/S100B(ββ) complex as a whole was stable, with the total Cα RMSD remaining around ~3 Å in all five independent 100-ns CHARRM production simulations (Fig. S2). All bound calcium ions remain intact throughout all simulations. However, the p53/S100B(ββ) binding interface were highly dynamic. As shown in Fig. 3A, the (self) Cα RMSD of the p53 peptide fluctuated between 2–3 Å in most simulations but could increase up to greater than 5 Å in some cases (3 out of the 10 independent interfaces). Large self RMSD values are strongly correlated with significant loss of helicity (e.g., see final snapshots shown in Fig. S1B). On the ensemble level (including the last 60 ns from all production trajectories), only the N-terminal region of p53 (residues 375–383) remains highly helical (with ~75% helicity), but the last helix turn (residues 384–387) is rarely formed (Fig. 3D). The pervious experimental assignment of the p53 helical segment (376-387) was based on Hα chemical shift indexing and medium range NOE connectivity[32]. It is probable that the current simulation underestimates the helicity in the p53-NRD C-terminus. On the other hand, absence of experimentally resolved $\alpha\beta_{i,i+3}$, $\alpha N_{i,i+3}$ medium range NOEs in the last helical turn appears to be consistent with at least reduced helicity in this region. Interestingly, the Cα binding RMSD of p53 (as defined in Methods) rapidly increased to around 4–8 Å in all simulations regardless of the starting conformations. Helix unfolding clearly contributes to the large p53 binding RMSD values observed; but large binding RMSD appears to more of a consequence of substantial changes in p53 binding mode.

As illustrated in Fig. 4, helix unwinding at the p53-NRD C-terminus allows key nonpolar residues, including Leu383 and Phe385, to penetrate deeper into the hydrophobic binding pocket of S100B(ββ) and make more extensive intermolecular contacts (see Fig. 2). Importantly, these hydrophobic interactions appear to play a key role in anchoring p53-NRD at the interface; for example, even though the C-terminal segment is less helical (Fig. 3D), it displays smaller fluctuations (Fig. 3C). The p53-NRD N-terminus appears to be stabilized mainly by interactions between Arg379 with several Glu residues on S100(ββ), including Glu45, Glu46, Glu49 and Glu51 on the hinge between the second and third helices (see Fig. 2 and Table S1). The presence of multiple complimentary charges on the hinge appears to facilitate the p53-NRD N-terminus to

sample alternative orientations in the bound state, leading to large fluctuations observed in simulations (Fig. 3C). Such conformational heterogeneity reduces the entropy penalty of p53 folding upon binding, and might help to explain the ability of p53-NRD to bind S100(ββ) with a dissociation constant $K_D \sim 23$ μM[44], despite having a small hydrophobic interface that is mainly anchored by a single hydrophobic residue (Leu383). Interestingly, although p53-NRD Lys386 was proposed to provide another anchoring point by interacting with S100(ββ) Glu86[32], it displays a strong tendency to remain solvated. In fact, solvation of Lys386 and deeper burial of hydrophobic side chains (particularly Leu383) appear to be two key driving forces that drive the observed changes in p53-NRD conformation.

The above qualitative observations are further supported by comparing the residue-residue contact maps calculated from the NMR and simulated ensembles. The maps are shown in Fig. 2, with all contacts observed with >0.2 probabilities listed in Table S1. The contact analysis clearly shows that the simulated ensemble predicts the p53 segment centered at Leu383 and Phe385 to be involved in more extensive contacts than in the NMR ensemble, while either the very N- or C-terminus (particularly Lys386) makes fewer stable contacts. Interestingly, the overall pattern of contacts calculated from the simulated ensemble appears more consistent with the set of experimentally resolved (weak) intermolecular NOEs (marked as red crosses in Fig. 2) than the one derived from the original NMR ensemble. For example, the NMR ensemble gives rise to a sparser contact map (Fig. 2A), with few or no residue-residue contacts in several regions where NOEs are experimentally observed (e.g., near S100B 44 – p53 385, S100B 55 – p53 383, and S100B 55 – p53 387). In contrast, the simulated ensemble yields significantly populated contacts in these regions (Fig. 2B). In addition, substantial conformational heterogeneity in the N-terminal segment of p53-NRD in the bound state could explain an apparent paucity of resolvable intermolecular NOEs for this region (only three very week ones were resolved; see Table 1). The consensus sequence of S100B-binding proteins have been suggested to contain a positive charge followed by several hydrophobic residues[44;45]; however, the exact sequence pattern appears to allow large variance, such as (K/R)(L/I)xWxxIL[45] or (K/R)(L/I)(P/S/N/D)(W/L/I)(S/D/L)x(L/I)(L/F)[44]. Such variance could suggest that conformational heterogeneity observed at the p53/S100B(ββ) interface might be present in complexes formed by other S100B-binding peptides.

Validation of the simulated ensemble: force field dependence and NOE violation analysis

Importantly, the observed "fuzziness" in the p53/S100B(ββ) interface does not seem to be an artifact of the CHARMM22/CMAP force field. Substantial instabilities have been observed consistently in previous simulations of the complex using other force fields[31;34;35]. For example, p53-NRD was found to be very flexible and to unfold significantly in the bound state during a 20-ns explicit solvent simulation in the Amber ff99 force field[31]. Allen *et al.* found that the C-terminus of p53-NRD unfolded during one of their 100-ns explicit solvent simulations in the GROMOS 53a6 force field[34]. We have further verified the force field dependence using an additional 100-ns simulation of the complex in the Amber ff99SB force field[43], which is arguably one of the best force fields optimized for simulating protein conformational equilibria[46]. The results, summarized in Fig. S3, are highly consistent with the observations from the CHARMM simulations. Specifically, the p53 helix was also observed to unfold substantially, especially at the C-terminus, to allow more extensive hydrophobic contacts with S100B(ββ) (Fig. S3B). Both N- and C-termini of p53-NRD remains highly dynamic in the S100B(ββ)-bound state, and substantial conformational heterogeneity exists at the interface. There is also a notable resemblance between the residue-residue contact maps shown in Fig. S3D (Amber) and Fig. 2B (CHARMM), despite the substantial difference in amount of sampling. The genuineness of the observed dynamic p53/S100B(ββ) interface is further supported by an independent simulation of the bound state using the OPLS-AA/L force field published while the current work was under review[47]. It was also observed that p53-NRD underwent substantial helix unfolding, particularly at the C-terminus, to allow more extensive contacts of the p53 segment centered at Leu383 and Phe385 with S100B((ββ).

As a more direct validation of the simulated ensemble, we back-calculated NOE-like distances between groups of atoms involved in all experimentally resolved intermolecular NOEs and compared them with the assigned upper bounds (see Methods). The results are summarized in Table 1, and histograms of NOE-like distances between pairs of atom groups involved in six selected intermolecular NOEs are shown in Fig. 5. The analysis shows that the simulated ensemble as a whole is highly consistent with the experimental NOE set, with only 6 of the 33 intermolecular NOEs violated by over 0.5 Å and only two violated by >2.0 Å. Both severely violated NOEs involve either Arg379 or Lys386, which, as discussed above, have a strong tendency to become solvent-exposed and likely drive the observed conformational changes (see Fig. 4). Importantly, substantial conformational dynamics and heterogeneity persisted throughout the simulations after initial relaxation, manifested as large fluctuation in NOE-like distances over

time (see Fig. S4) and divergence in the histograms of NOE-like distances calculated from different trajectories (Fig. 5). It should be emphasized that consistency with the experimental NOE restraints, although an important validation, is not sufficient to establish the correctness of the entire heterogeneous ensemble due to underdetermination[23]. In addition, because of the r^{-6} averaging nature of NOE, calculated (or measured) ensemble-averaged NOE distances are not sensitive to conformations with much larger distances within the ensemble.

Clustering analysis: existence of multiple conformational sub-states

The average conformations of 8 largest clusters of the simulated ensemble identified by *k*-mean clustering using p53 binding C_α RMSD are shown in Fig. 6. These average structures further illustrate the broad manifold of conformational sub-states sampled by p53-NRD in the bound state. The p53 conformation itself appears to be somewhat similar in these clusters, with a rather stably formed helix at the N-terminus followed by a coiled C-terminal tail; however, substantial differences exist in the orientation of p53-NRD and specific interactions that it forms with S100B(ββ). The p53 peptide appears to pivot around Leu383, which forms many contacts with residues in the hydrophobic binding pocket of S100B(ββ) (see Fig. 2 and Table S1). The more helical N-terminus, stabilized mainly by electrostatic interactions, however, can sample many alternative orientations as represented by these clusters.

We have further performed NOE violation analysis on sub-ensembles of structures that belong to the 10 most populated clusters. The results are summarized in Fig. 7. Several interesting observations can be made. First, all intermolecular NOEs are satisfied by two or more of the 10 largest clusters, including those that are (severely) violated on the whole ensemble level (e.g., R379-K48 and K386-L44; see Table 1). Second, only a fraction of NOEs (6 out of 31) is consistently satisfied by all clusters analyzed, and these NOEs mainly involve Leu383 and the neighboring Met384. This is consistent with the structural analysis showing that the segment around Leu383 anchors the p53/S100B(ββ) interface. The third and arguably the most interesting observation is that most intermolecular NOEs are satisfied only by a few clusters. This result illustrates one of the major challenges in structural characterization of dynamic protein states, including both unbound IDPs and fuzzy complexes: That structural restraints derived from ensemble methods such as NMR or FRET often reflect conformational features that do not necessarily co-exist in the same conformational sub-state. This, again, highlights the importance of properly recognizing and representing fuzzy complexes in structural studies of IDP interactions.

Access to p53 PTM sites: inhibition of p53 activity

To examine how effectively the simulated heterogeneous ensemble can hinder access to PTM sites in p53-NRD, we analyzed the average buried surface areas of p53 residues. The results, shown in Fig. 8, demonstrate that the profiles of residue buried surface areas from the NMR and simulated ensembles are very similar. Key PTM sites within p53-NRD include two phosphorylation sites (Ser376 and Thr377) and two acetylation sites (Arg379 and Lys386). Clearly, all these PTM sites except Thr377 remain sterically hindered despite the dynamic nature of p53-NRD in the simulated bound state. An important difference is that Thr377 becomes fully solvent-exposed and could potentially be phosphorylated by protein Kinase C (PKC); however, the helical state of the p53-NRD N-terminal segment, coupled with steric shielding of the neighboring residues, is likely sufficient to prevent access of Thr377 to the PKC active site for efficient phosphorylation. In addition, previous in vitro experiments have showed that either S376A or T377A alone could reduce p53 activity to a level that might be achieved by co-transfection of S100B(ββ) (~50%)[48]. Therefore, the proposed heterogeneous ensemble could fully support S100(ββ) inhibition of p53 activity.

Discussion

Recent recognition of the prevalence of intrinsic disorder in biology has drastically expanded the perception of how protein structure may mediate function. Substantial progress has been made in understanding the structural basis and molecular mechanisms of IDP function, but mainly in the context of coupled binding and folding that involve well-structured bound states[10;12;13]. Examples have also emerged to suggest the generality and importance of structural disorder (or "fuzziness") in the bound state of IDPs[14;15]. This is an important recognition that requires new ways of thinking about and studying IDP interaction and regulation. In particular, conformational ensembles are necessary for describing dynamic bound states, and conventional restrained molecular dynamics-based structural calculation protocols are no longer applicable. However, important challenges exist in calculating heterogeneous structure ensembles, and this is mainly due to the fundamental limitation that ensemble-averaged properties alone do not provide sufficient constraint (information) to uniquely define representative structure ensembles for heterogeneous protein states[16;21]. Instead, with important advances in protein force fields and sampling methodologies in recent years, physics-based atomistic simulations may provide an effective alternative means to generate *de novo* structural ensembles of disordered proteins states[49].

For the case of the p53/S100B(ββ) complex, analysis of the NMR data and calculated structure ensemble suggests that the p53/S100B(ββ) interface is highly dynamic. Extensive atomistic explicit solvent simulations were performed using two state-of-the-art empirical protein force fields, which consistently revealed substantial heterogeneity in the p53-NRD conformation and how it binds with S100B(ββ). Specifically, helix unwinding at the p53-NRD C-terminus allows key hydrophobic residues (Leu383 and Phe385) to make more extensive contacts with S100B(ββ); the N-terminal segment, although highly helical, displays great flexibility in its packing with S100B(ββ). Additional NMR experiments, particularly relaxation measurements, will be necessary to provide direct experimental evidence on the dynamic nature of the S100B(ββ)-bound state of p53-NRD. Nonetheless, the simulated dynamic ensemble is highly consistent with experimental intermolecular NOE restraints, and provides similar shielding of key PTM sites in p53-NRD to support S100B(ββ) inhibition of p53 activity. The simulation thus offers a plausible alternative structural interpretation of how S100B(ββ) regulate p53 activity, and provides new insights for design of rational strategies for modulating this interaction for cancer therapy. The current study also demonstrates the importance and efficacy of an integrated computational and experimental strategy for understanding fuzzy complexes.

Methods

Explicit solvent simulations

The NMR ensemble (PDB: 1dt7[32]) consists of 40 structures and each contains two p53-NRD (residues 369-388) in complex with a Ca^{2+}-loaded S100B(ββ) dimer. Five diverse conformers were first selected based on mutual backbone root-mean-square distances (RMSDs). The selected models were models 1, 21, 23, 30 and 31 from the NMR ensemble (see Fig. S1). These structures were then used to generate the initial configurations for independent simulations to facilitate conformational sampling. The NMR study showed that the N-terminal segment of p53-NRD (residues 369–373) remained disordered and had little interaction with S100B(ββ); therefore, only residues 374–388 of p53 (sequence: GQ STSRH KKLMF KTE) were included in simulations. Removing the extended C-terminal tail also allowed ~25% smaller water boxes to be used. The truncated complex (including all four bound calcium ions) was solvated in TIP3P explicit water. Nine free calcium ions were added to neutralize the total charge. The final solvated system contains ~25,000 atoms in a periodic rectangular box of ~73 Å × 63 Å × 55 Å.

All five solvated initial conformations were equilibrated in CHARMM[50;51] using multiple stages of energy minimization and constant pressure/temperature (NPT) simulation with gradually reducing harmonic positional restraints on the complex. The CHARMM22/CMAP force field[40-42] was used. After equilibration, five 100-ns NPT production simulations were performed at 300 K and under atmospheric pressure. To prevent global tumbling of the complex (such that a rectangular periodic box could be used), weak harmonic positional restraints with a force constant of 0.1 kcal/mol/Å^2 were placed on the backbone heavy atoms of S100B(ββ) that were at least 10 Å away from any p53 atoms. These harmonic restraints are not expected to affect sampling of p53 conformation and interaction, because the S100B(ββ) dimer should remain stably folded with or without bound p53[32]. SHAKE[52] was applied to constrain the lengths of all bonds involving hydrogen atoms, and a 2.0-fs dynamic time step was used. Particle Mesh Ewald (PME) was used to treat the long-range electrostatic interactoins[53], and the van der Waals interactions were smoothly switched off from 12 to 13 Å. Snapshots were saved every 2 ps during the production simulations. Because there are two independent p53/S100B(ββ) interfaces in the system, five 100-ns simulations together yield a total of 1.0 μs effective sampling of the p53/S100B(ββ) interaction.

To examine the force field dependence, the minimized structure model 1 of the NMR ensemble was also used to initiate an independent simulation in the Amber ff99SB force field[43], performed using GROMACS[54]. The protein complex was solvated in a TIP3P cubic box with sides of 70 Å. Counter ions were added to the system using *genion* program include with GROMACS program suite. All bonds involving hydrogen atoms were restrained using the SHAKE. PME was used to treat long-range electrostatics, and van de Waals interactions were smoothly switched off at 12 Å. The system was energy minimized, followed by a 20-ns equilibration with a 2-fs time-step. The production simulation was performed for 100 ns. The simulation trajectory was processed to remove periodic boundary wrapping and converted to the CHARMM format for analysis.

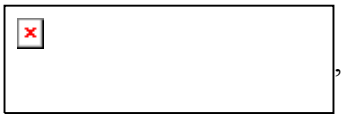
Structural and Clustering Analysis

All structural and clustering analysis were performed using CHARMM, the Multiscale Modeling Tools for Structural Biology (MMTSB) toolset[55], and in-house scripts. Molecular visualization was generated using VMD[56]. Residue helicity was calculated using the COOR SECS module. Evolution of RMSD values of various segments, together with helicity analysis, suggests that the complex reached some type of (local) equilibrium after 20-40 ns (e.g., see Fig. S2 and Fig. 3). As such, conformations sampled during the last 60 ns of all five CHARMM trajectories were

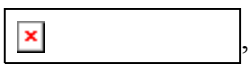
included in the simulated conformational ensembles of the complex. All clustering and NOE analysis were performed on this ensemble unless otherwise noted. For clustering, the simulated ensemble was first under-sampled by only including snapshots every 100 ps during the last 60 ns of each production trajectory. Each snapshot was then split to generate two independent monomeric p53/S100B(ββ) complex structures. The resulting 6000-member ensemble was clustered using the fixed radius clustering algorithm as implemented in the MMTSB/enscluster.pl tool (with –kclust option). The distance between conformers for clustering was provided by p53 C_α RMSD, which was calculated by first aligning the whole monomeric complex using the backbone atoms of S100B(ββ). The RMSD values as calculated reflect differences in both internal conformation and the binding pose of p53-NRD, and will be referred to as "binding RMSD" for the rest of the text. The cutoff radius used in *k*-mean clustering was empirically selected to be 5 Å. Further NOE and helicity analysis was performed for each of the 10 most populated clusters, which presumably represent key sub-states sampled by p53 when bound to S100B(ββ). It is important to emphasize that conformational sampling achieved in multiple 100-ns explicit solvent simulations remains limited and that the final ensemble as constructed without reweighing independent trajectories is not a rigorous Boltzmann ensemble; therefore, all calculated conformational properties should be considered semi-quantitative, if not only qualitative, estimates.

NOE Violation Analysis

The simulated ensemble was validated mainly through NOE violation analysis. For this, the original NMR restraint data files were downloaded from the direct link provided with the Protein Data Bank entry 1dt7, which included 33 unique intermolecular NOEs for each p53 peptide (see Table 1). Coordinates of all the atoms involved in each NOE assignment were extracted from the production trajectories. At each time point, NOE-like distances between two selected atom groups were calculated using the r^{-6} summation scheme[57],



where r_{ij} is the distance between atoms i from one atom selection and j from the other as specified in the experimental NOE restraint set. The ensemble (time) averaged NOE-like distances were then calculated as:



where <> represent averaging over all (time) snapshots included in a particular ensemble. Note that the r^{-6}-weighted nature of NOE averaging renders r_{NOE} highly sensitive to small population of conformations with small $r_{NOE}(t)$ and at the same time rather insensitive to fraction of populations with significantly larger $r_{NOE}(t)$. We also note that protein internal dynamics can modulate the cross relaxation rates and can be included in NOE analysis of continuous MD trajectories[43;58]. The effects of angular fluctuation are ignored in the above analysis due to the need to include conformations sampled from multiple independent simulations. Nonetheless, comparative analysis using the model 1 trajectory suggests that neglecting ps-ns timescale motions in NOE analysis only leads to limited under-estimation of the back-calculated NOE distances (see Table S3 and the following text).

Acknowledgements

The authors thank Debabani Ganguly and Weihong Zhang for helpful discussions and Elizabeth Ploetz for assistance with GROMACS, This work was supported by the National Science Foundation (MCB 0952514) and the KSU Johnson Center for Basic Cancer Research. This work is contribution number 12-402-J from the Kansas Agricultural Experiment Station.

References

1. Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* **293**, 321-331.
2. Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradovic, Z. & Dunker, A. K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. *Journal of Molecular Biology* **323**, 573-584.
3. Tompa, P. (2002). Intrinsically unstructured proteins. *Trends in Biochemical Sciences* **27**, 527-533.
4. Romero, P., Obradovic, Z., Li, X. H., Garner, E. C., Brown, C. J. & Dunker, A. K. (2001). Sequence complexity of disordered protein. *Proteins-Structure Function and Genetics* **42**, 38-48.
5. Zhang, C. & Ma, J. (2012). Folding helical proteins in explicit solvent using dihedral-biased tempering. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8139-44.
6. Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D. & Nussinov, R. (2003). Extended disordered proteins: targeting function with less scaffold. *Trends in Biochemical Sciences* **28**, 81-85.
7. Tompa, P., Szasz, C. & Buday, L. (2005). Structural disorder throws new light on moonlighting. *Trends in Biochemical Sciences* **30**, 484-489.
8. Hilser, V. J. & Thompson, E. B. (2007). Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8311-8315.

9. Oldfield, C. J., Meng, J., Yang, J. Y., Yang, M. Q., Uversky, V. N. & Dunker, A. K. (2008). Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics* **9 Suppl 1**, S1.
10. Wright, P. E. & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology* **19**, 31-38.
11. Best, R. B. (2012). Atomistic molecular simulations of protein folding. *Current Opinion in Structural Biology* **22**, 52-61.
12. Click, T. H., Ganguly, D. & Chen, J. H. (2010). Intrinsically Disordered Proteins in a Physics-Based World. *International Journal of Molecular Sciences* **11**, 5293-5309.
13. Chen, J. H. (2012). Towards the physical basis of how intrinsic disorder mediates protein function. *Archives of Biochemistry and Biophysics* **524**, 123-131.
14. Tompa, P. & Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends in Biochemical Sciences* **33**, 2-8.
15. Ganguly, D., Otieno, S., Waddell, B., Iconaru, L., Kriwacki, R. W. & Chen, J. (2012). Electrostatically Accelerated Coupled Binding and Folding of Intrinsically Disordered Proteins. *Journal of Molecular Biology*, (in press).
16. Mittag, T., Kay, L. E. & Forman-Kay, J. D. (2010). Protein dynamics and conformational disorder in molecular recognition. *Journal of molecular recognition* **23**, 105-16.
17. Graham, T. A., Ferkey, D. M., Mao, F., Kimelman, D. & Xu, W. (2001). Tcf4 can specifically recognize beta-catenin using alternative conformations. *Nature Structural Biology* **8**, 1048-52.
18. Sigalov, A., Aivazian, D. & Stern, L. (2004). Homooligomerization of the cytoplasmic domain of the T cell receptor zeta chain and of other proteins containing the immunoreceptor tyrosine-based activation motif. *Biochemistry* **43**, 2049-61.
19. Mittag, T., Orlicky, S., Choy, W. Y., Tang, X. J., Lin, H., Sicheri, F., Kay, L. E., Tyers, M. & Forman-Kay, J. D. (2008). Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17772-17777.
20. Mittag, T., Marsh, J., Grishaev, A., Orlicky, S., Lin, H., Sicheri, F., Tyers, M. & Forman-Kay, J. D. (2010). Structure/Function Implications in a Dynamic Complex of the Intrinsically Disordered Sic1 with the Cdc4 Subunit of an SCF Ubiquitin Ligase. *Structure* **18**, 494-506.
21. Fisher, C. K. & Stultz, C. M. (2011). Constructing ensembles for intrinsically disordered proteins. *Current Opinion in Structural Biology* **21**, 426-31.
22. Marsh, J. A. & Forman-Kay, J. D. (2012). Ensemble modeling of protein disordered states: Experimental restraint contributions and validation. *Proteins-Structure Function and Bioinformatics* **80**, 556-572.
23. Ganguly, D. & Chen, J. (2009). Structural interpretation of paramagnetic relaxation enhancement-derived distances for disordered protein states. *Journal of Molecular Biology* **390**, 467-77.
24. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. (2003). The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance* **160**, 65-73.
25. Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. (1991). p53 mutations in human cancers. *Science* **253**, 49-53.
26. Soussi, T., Ishioka, C., Claustres, M. & Beroud, C. (2006). Locus-specific mutation databases: pitfalls and good practice based on the p53 experience. *Nature reviews. Cancer* **6**, 83-90.

27. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. (2008). Intrinsically disordered proteins in human diseases: Introducing the D-2 concept. *Annual Review of Biophysics* **37**, 215-246.
28. Lin, J., Blake, M., Tang, C., Zimmer, D., Rustandi, R. R., Weber, D. J. & Carrier, F. (2001). Inhibition of p53 transcriptional activity by the S100B calcium-binding protein. *The Journal of biological chemistry* **276**, 35037-41.
29. Markowitz, J., MacKerell, A. D., Carrier, F., Charpentier, T. H. & Weber, D. J. (2005). Design of inhibitors for S100B. *Current Topics in Medicinal Chemistry* **5**, 1093-1108.
30. Markowitz, J., MacKerell, A. D. & Weber, D. J. (2007). A search for inhibitors of S100B, a member of the S100 family of calcium-binding proteins. *Mini-Reviews in Medicinal Chemistry* **7**, 609-616.
31. Whitlow, J. L., Varughese, J. F., Zhou, Z. G., Bartolotti, L. J. & Li, Y. M. (2009). Computational screening and design of S100B ligand to block S100B-p53 interaction. *Journal of Molecular Graphics & Modelling* **27**, 969-977.
32. Rustandi, R. R., Baldisseri, D. M. & Weber, D. J. (2000). Structure of the negative regulatory domain of p53 bound to S100B(beta beta). *Nature Structural Biology* **7**, 570-574.
33. Luo, R., David, L., Hung, H., Devaney, J. & Gilson, M. K. (1999). Strength of solvent-exposed salt-bridges. *Journal of Physical Chemistry B* **103**, 727-736.
34. Allen, W. J., Capelluto, D. G., Finkielstein, C. V. & Bevan, D. R. (2010). Modeling the relationship between the p53 C-terminal domain and its binding partners using molecular dynamics. *The journal of physical chemistry. B* **114**, 13201-13.
35. Chen, J. H. (2009). Intrinsically disordered p53 extreme C-terminus binds to S100B(betabeta) through "fly-casting". *Journal of the American Chemical Society* **131**, 2088-9.
36. Wang, J., Cieplak, P. & Kollman, P. A. (2000). How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* **21**, 1049-1074.
37. Oostenbrink, C., Villa, A., Mark, A. E. & van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry* **25**, 1656-76.
38. Chen, J. H., Im, W. P. & Brooks, C. L. (2006). Balancing solvation and intramolecular interactions: Toward a consistent generalized born force field. *Journal of the American Chemical Society* **128**, 3728-3736.
39. Im, W. P., Lee, M. S. & Brooks, C. L. (2003). Generalized born model with a simple smoothing function. *Journal of Computational Chemistry* **24**, 1691-1702.
40. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* **102**, 3586-3616.
41. Mackerell, A. D., Feig, M. & Brooks, C. L. (2004). 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. *Journal of Computational Chemistry* **25**, 1400-1415.
42. MacKerell, A. D., Feig, M. & Brooks, C. L. (2004). Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society* **126**, 698-699.

43. Brueschweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M. & Ernst, R. R. (1992). Influence of rapid intramolecular motion on NMR cross-relaxation rates. A molecular dynamics study of antamanide in solution. *Journal of the American Chemical Society* **114**, 2289-2302.
44. Wilder, P. T., Lin, J., Bair, C. L., Charpentier, T. H., Yang, D., Liriano, M., Varney, K. M., Lee, A., Oppenheim, A. B., Adhya, S., Carrier, F. & Weber, D. J. (2006). Recognition of the tumor suppressor protein p53 and other protein targets by the calcium-binding protein S100B. *Biochimica et Biophysica Acta* **1763**, 1284-97.
45. Ivanenkov, V. V., Jamieson, G. A., Jr., Gruenstein, E. & Dimlich, R. V. (1995). Characterization of S-100b binding epitopes. Identification of a novel target, the actin capping protein, CapZ. *Journal of biological chemistry* **270**, 14651-8.
46. Wickstrom, L., Okur, A. & Simmerling, C. (2009). Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophysical Journal* **97**, 853-6.
47. Staneva, I., Huang, Y., Liu, Z. & Wallin, S. (2012). Binding of Two Intrinsically Disordered Peptides to a Multi-Specific Protein: A Combined Monte Carlo and Molecular Dynamics Study. *Plos Computational Biology* **8**, e1002682.
48. Youmell, M., Park, S. J., Basu, S. & Price, B. D. (1998). Regulation of the p53 protein by protein kinase C alpha and protein kinase C zeta. *Biochemical and Biophysical Research Communications* **245**, 514-8.
49. Chen, J. (2012). Towards the physical basis of how intrinsic disorder mediates protein function. *Archives of Biochemistry and Biophysics* **524**, 123-31.
50. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry* **4**, 187-217.
51. Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M. & Karplus, M. (2009). CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry* **30**, 1545-1614.
52. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics* **23**, 327-341.
53. Darden, T., York, D. & Pedersen, L. (1993). Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **98**, 10089-10092.
54. Hess, B., Kutzner, C., vanderSpoel, D. & Lindahl, E. (2008). GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4**, 435-447.
55. Feig, M., Karanicolas, J. & Brooks, C. L. (2004). MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *Journal of Molecular Graphics & Modelling* **22**, 377-395.
56. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33-&.
57. Fletcher, C. M., Jones, D. N., Diamond, R. & Neuhaus, D. (1996). Treatment of NOE constraints involving equivalent or nonstereoassigned protons in calculations of biomacromolecular structures. *Journal of Biomolecular NMR* **8**, 292-310.
58. Olejniczak, E. T., Dobson, C. M., Karplus, M. & Levy, R. M. (1984). Motional averaging of proton nuclear Overhauser effects in proteins. Predictions from a molecular dynamics simulation of lysozyme. *Journal of the American Chemical Society* **106**, 1923-1930.

Table 1. Summary of NOE violation analysis of the simulated ensemble.

Atom Pairs	R_{nmr} (Å) [a]	R_{sim} (Å) [a]	Violation (Å)
L383 HD* - L3 HD* [b]	6.0	6.2	0.2
L383 HD* - M7 HE* [b]	5.0	6.2	1.2
Q375 HB* - E51 HG*	6.0	7.0	1.0
R379 HB* - K48 HG*	5.0	7.0	2.0
R379 HD* - V52 HG2*	6.0	3.0	–
L383 HD* - V52 HG2*	6.0	3.1	–
L383 HD* - V56 HG*	5.0	2.2	–
L383 HD* - V77 HG2*	6.0	7.3	1.3
L383 HD* - M79 HE*	3.3	3.1	–
L383 HB* - M79 HE*	5.0	4.0	–
L383 HD* - V80 HG2*	4.2	2.9	–
L383 HD* - V80 HG1*	5.0	3.4	–
L383 HD* - T82 HG2*	6.0	7.1	1.1
L383 HB* - A83 HB*	5.0	3.1	–
L383 HD* - A83 HB*	5.0	4.7	–
M384 HB* - K55 HG*	6.0	4.5	–
M384 HG* - V56 HG*	6.0	2.3	–
M384 HE* - V56 HG*	6.0	2.8	–
M384 HE* - L44 HD*	6.0	3.4	–
F385 HE* - M79 HE*	6.0	4.9	–
F385 HZ - M79 HE*	6.0	6.5	0.5
F385 HE* - V80 HG2*	6.0	3.0	–
F385 HZ - V80 HG2*	6.0	4.3	–
F385 HD* - V80 HG1*	6.0	5.0	–
F385 HE* - A83 HB*	5.0	4.1	–
F385 HZ - A83 HB*	6.0	3.3	–
F385 HZ - L44 HD*	6.0	3.0	–
K386 HD* - T82 HG2*	5.0	3.4	–
K386 HG* - T82 HG2*	6.0	4.4	–
K386 HD* - L44 HD*	6.0	8.4	2.4
T387 HG* - K55 HG*	5.0	4.2	–
E388 HG* - M79 HE*	5.0	4.6	–
E388 HG* - T82 HG2*	6.0	4.2	–

[a] R_{nmr}: experimentally assigned upper bound; R_{sim}: ensemble-averaged NOE-like distance calculated from the simulated ensemble.

[b] Involving p53-S100B(ββ) side chains across the symmetric dimer interface.

Figure Legends

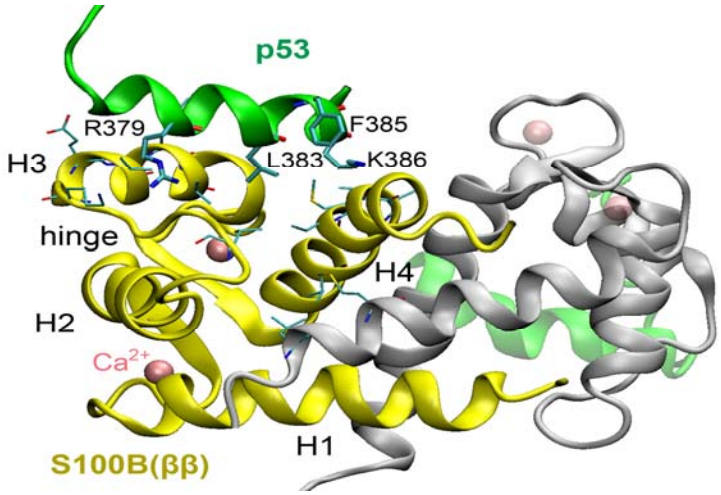


Figure 1. Cartoon representation of the p53/S100B(ββ) complex structure ((PDB:1dt7, model 1). Only one monomer of the dimeric complex is highlighted in solid colors. Bound calcium ions are displayed as pink spheres. Key residues of p53 are shown in thick sticks and labeled. All S100B(ββ) residues involved in NOE contacts with the p53 peptide highlighted are shown in thin sticks. Note that two of the intermolecular NOEs involve S100B(ββ) side chains across the symmetric dimer interface (Leu3 and Met7; see Table 1). Both residues locate deep at the bottom of the hydrophobic pocket where p53 binds. Four helices (H1-4) of S100B(ββ), together with the hinge between H2 and H3 are also labeled.

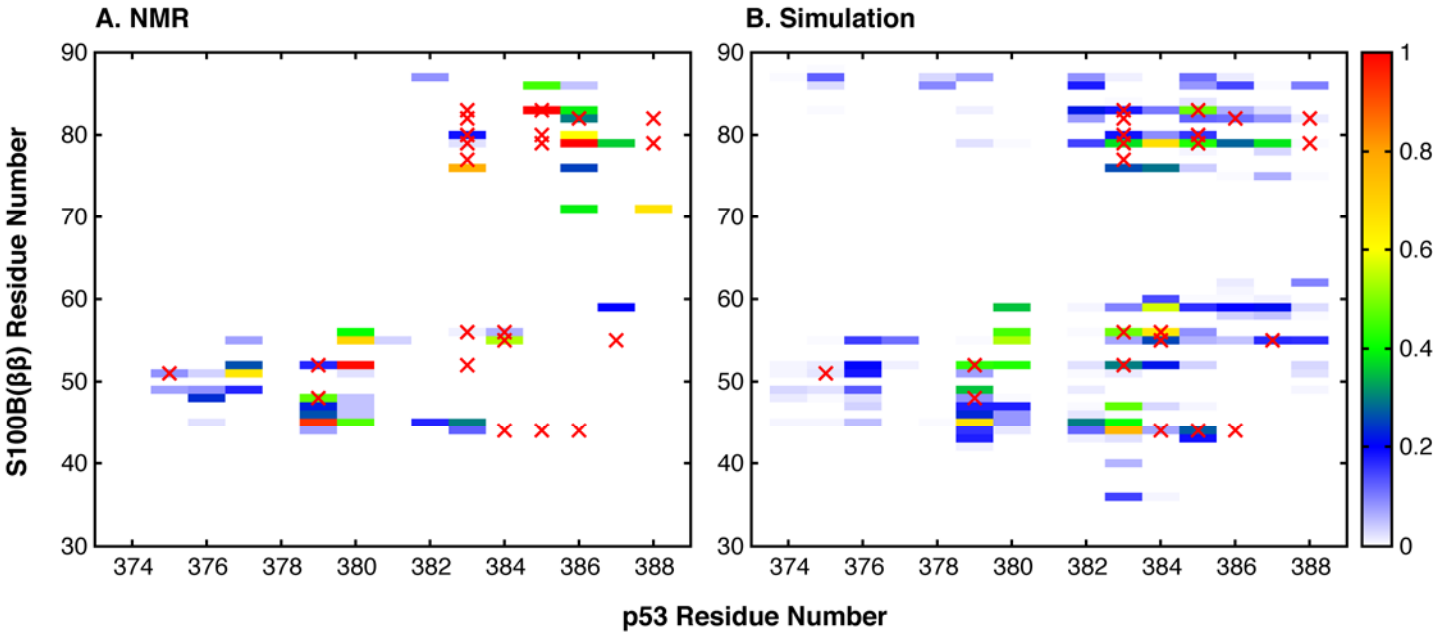


Figure 2. Intermolecular residue contact maps computed from A) the NMR ensemble (PDB: 1dt7) and B) the simulated ensemble. A contact is considered formed when the shortest distance between heavy atoms of two residues is no greater than 4.2 Å. The contacts are color-coded according to their probabilities. The red crosses mark the residue pairs that are involved in the experimentally resolved intermolecular NOEs (also see Table 1).

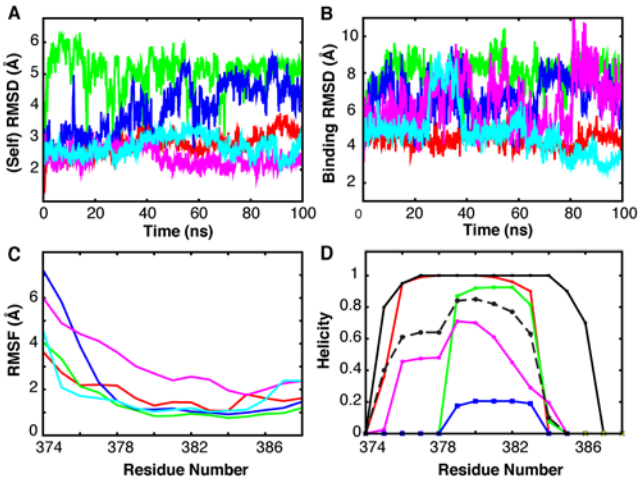


Figure 3. Conformational properties of p53 in the p53/S100B(ββ) complex. A) Cα (self) RMSD, B) binding RMSD and C) root mean square fluctuation (RMSF) of p53 (first monomer in the complex) calculated from five independent 100-ns simulations, initiated from NMR ensemble model 1 (red), 21 (green), 23 (blue), 30 (pink), and 31 (cyan), respectively. The RMSF profiles

were calculated by first aligning all snapshots using all Cα atoms. D) Averaged residue helicity profiles calculated from the NMR (solid black) and simulated ensembles (dashed black). A few representative helicity profiles for individual p53 monomers in various production simulations are also shown: red (first monomer from model 1 trajectory), green and pink (p53 monomers from model 21 trajectory), and blue trace (second p53 monomer from model 23 trajectory). Note that most other monomers yield helicity profiles similar to the red trace.

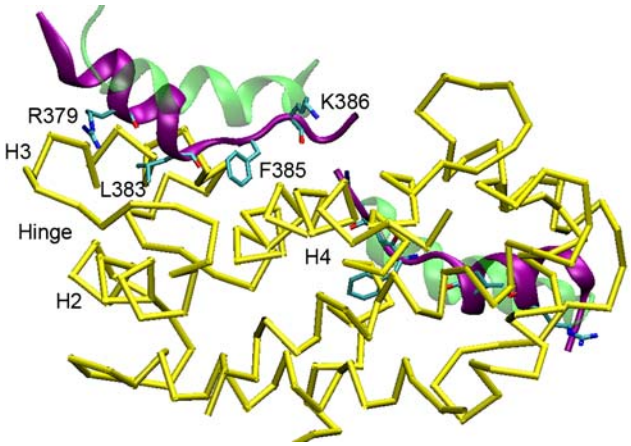


Figure 4. Two representative simulated conformations of p53 bound to S100(ββ). The two p53 monomers in the final snapshot of the 100-ns CHARMM simulation initiated from NMR ensemble model 1 are used (also see Fig. S1). S100B(ββ) is shown in yellow trace and p53 peptides in purple cartoon. Several key p53 residues are shown in sticks. The initial p53 conformations are also plotted in transparent green cartoon for reference.

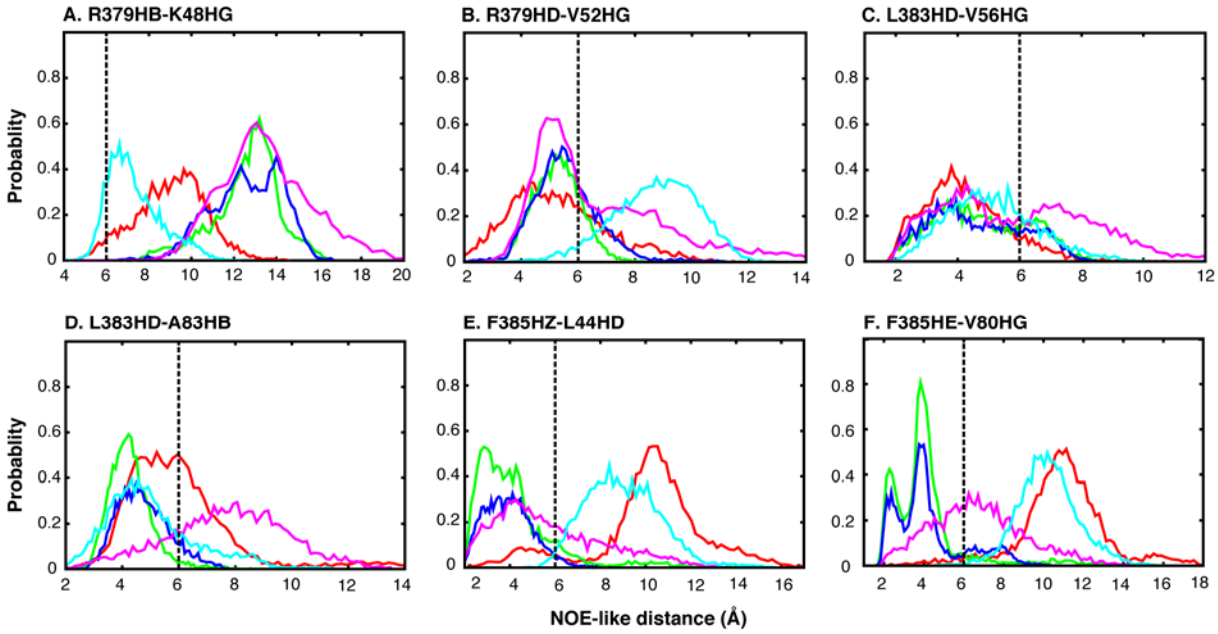


Figure 5. Histograms of NOE-like distances between pairs of atom groups involved in six selected intermolecular NOEs. All selected NOEs involve at least one of the key p53 residues for binding S100B(ββ), namely, Arg379, Leu383, and Phe385. These histograms were calculated from the last 60 ns of five independent simulations initiated from selected NMR models. The color-coding is the same as in Fig. 3A. The black dashed lines mark the experimentally assigned NOE upper bounds. Only the R379HB-K48HG NOE is violated (by 2.0 Å) when all five simulations are included (see Table 1).

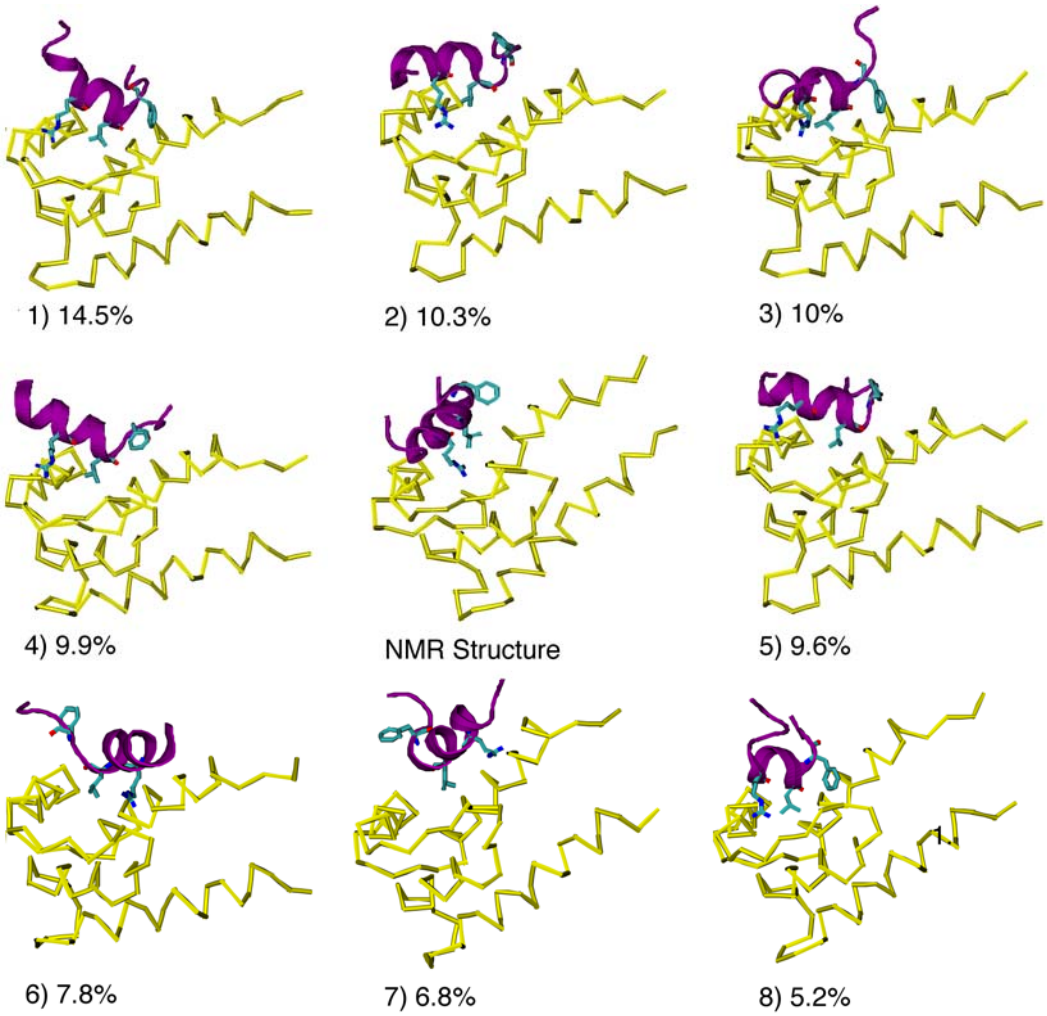


Figure 6. Average structures and populations of the eight largest clusters of the simulated ensemble. S100B(ββ) is shown in yellow trace and the p53 peptide in purple cartoon. Several key p53 residues are shown in sticks. The NMR structure (model 1) is shown in the middle for reference. All conformations are aligned using all Cα atoms of S100B(ββ).

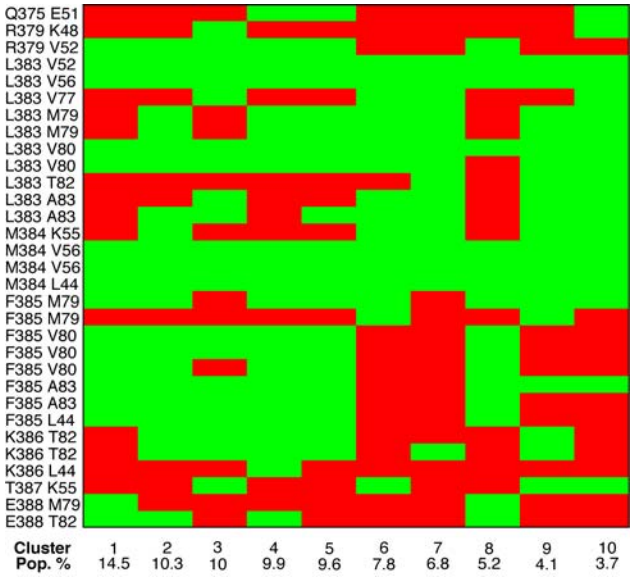


Figure 7. Summary NOE violation analysis of the top ten most populated clusters of the simulated ensemble. The identities of intermolecular NOEs are labeled on the y-axis. These NOEs are arranged in the same order as in Table 1. The analysis does not include the two NOEs that involve residues across the symmetric dimer interface. Red blocks mark NOEs that are violated by at least1.0 Å by individual clusters.

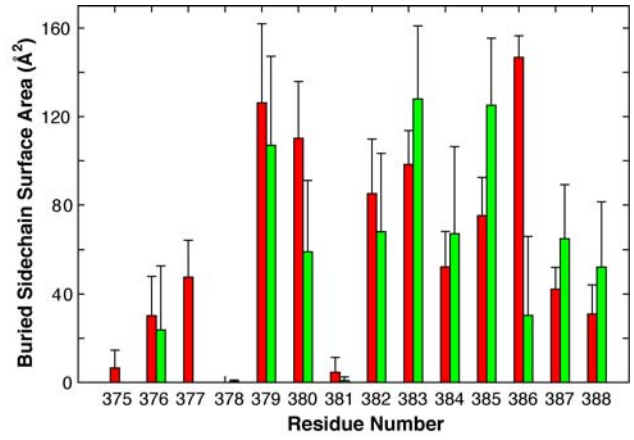


Figure 8: Average buried (solvent-accessible) surface areas of p53 residues from the NMR (red bars) and simulated (green bars) ensembles. The errors bars shown are standard deviations calculated from either the 40 members of the NMR ensemble or the five independent simulations.

Supporting Information

C. McDowell *et al.*

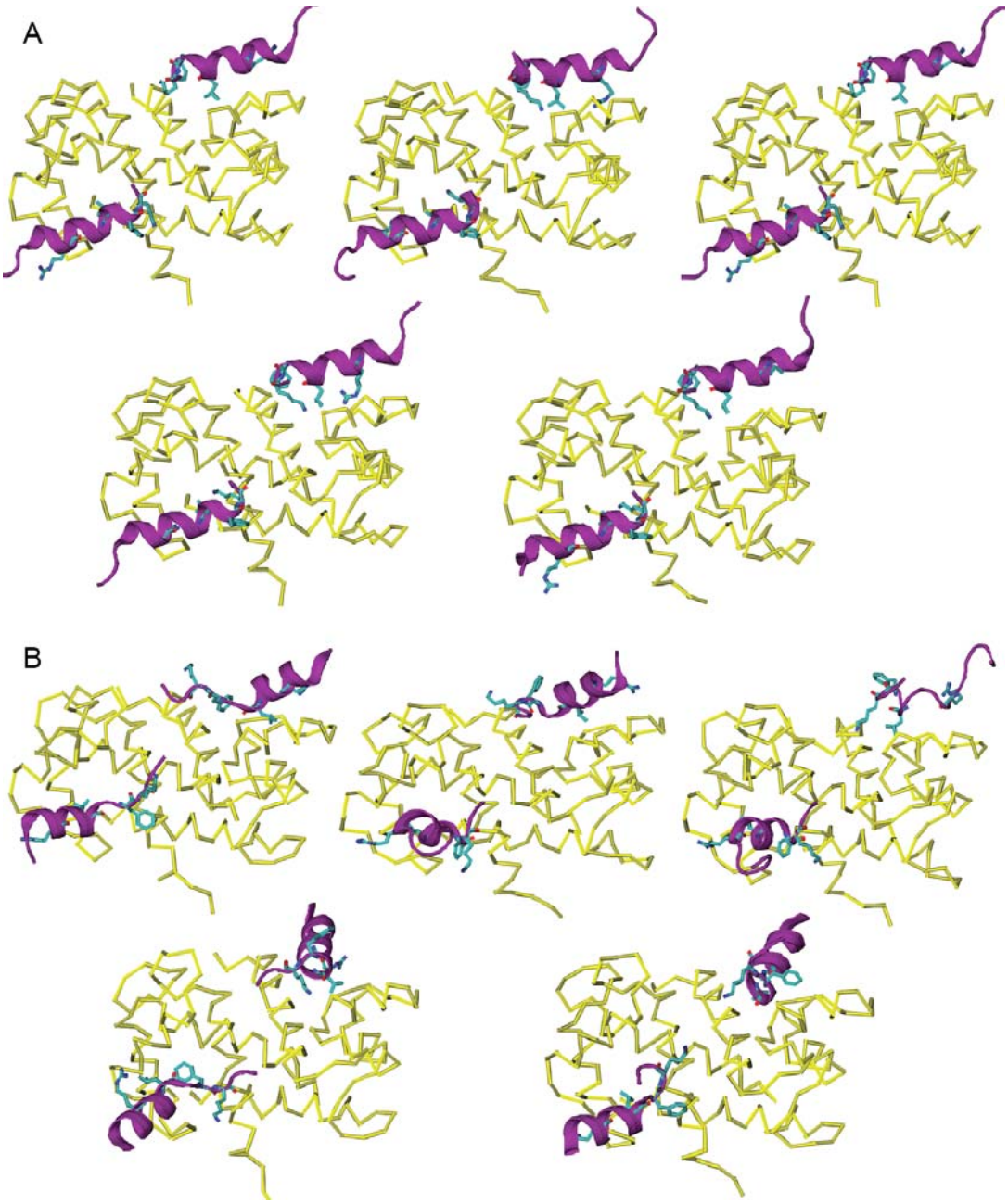


Figure S1. A) Five diverse models selected from the NMR ensemble (PDB: 1dt7). From top left to right, the model numbers are 1st, 21nd, 23rd, 30th, and 31st. B) Final snapshots after 100 ns production simulation in the CHARMM22/CMAP force field. S100B((ββ) is shown in yellow trace and p53 peptides in purple cartoon representations. Side chains of key p53 residues (R379, L383, and F385) are shown in sticks. All initial and final snapshots are displayed in the same orientation (by aligning the Cα atoms of S100B). Waters and ions are not shown for clarity.

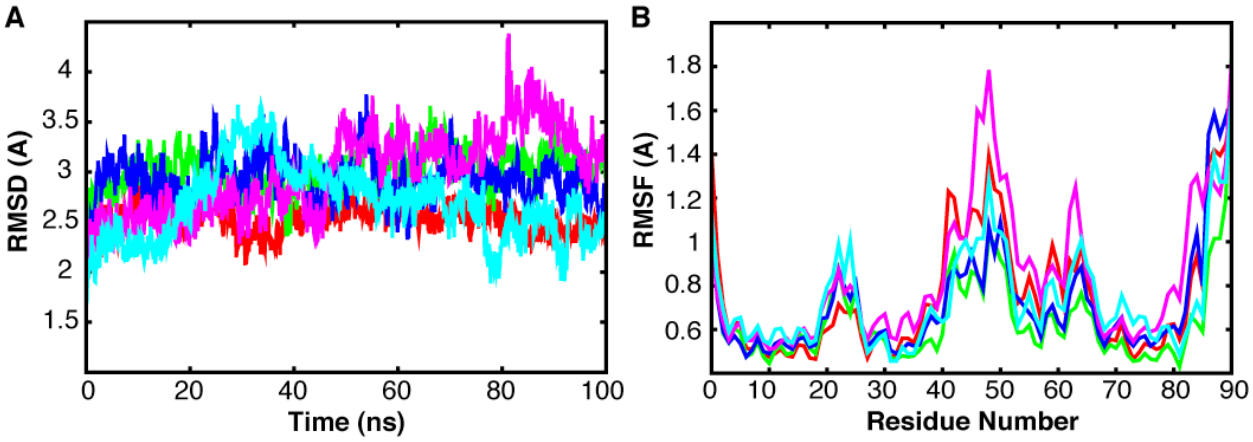


Figure S2. A) Cα RMSD of the whole complex from their corresponding initial conformations during five 100-ns production simulations, initiated from NMR ensemble model 1 (red), 21 (green), 23 (blue), 30 (pink), and 31 (cyan), respectively. B) Cα RMSF of S100B(ββ) (first monomer of the dimer) calculated from the last 60 ns of the production simulations. The curves are color coded in the same fashion as in panel A.

Table S1. Intermolecular contacts observed with >0.2 probabilities from the NMR and simulated ensembles.

NMR (PDB: 1dt7)			Simulation		
Residue Pairs		**Probabilities**	**Residue Pairs**		**Probabilities**
Ser 376	Lys 48	0.24	Ser 376	Val 52	0.20
Thr 377	Glu 51	0.65	Arg 379	Glu 45	0.67
Thr 377	Val 52	0.26	Arg 379	Glu 46	0.23
Arg 379	Glu 45	0.94	Arg 379	Glu 49	0.35
Arg 379	Glu 46	0.26	Arg 379	Val 52	0.44
Arg 379	Ile 47	0.23	His 380	Val 52	0.42
Arg 379	Lys 48	0.48	His 380	Lys 55	0.54
His 380	Glu 45	0.46	His 380	Val 56	0.46
His 380	Val 52	0.99	His 380	Thr 59	0.35
His 380	Lys 55	0.69	Lys 382	Glu 45	0.30
His 380	Val 56	0.40	Lys 382	Ala 83	0.22
Leu 383	Glu 45	0.30	Leu 383	Leu 44	0.76
Leu 383	Phe 76	0.78	Leu 383	Glu 45	0.40
Met 384	Lys 55	0.54	Leu 383	Ile 47	0.47
Phe 385	Ala 83	1.00	Leu 383	Val 52	0.30
Phe 385	Glu 86	0.45	Leu 383	Val 56	0.49
Lys 386	Gln 71	0.39	Leu 383	Phe 76	0.26
Lys 386	Phe 76	0.25	Leu 383	Met 79	0.38
Lys 386	Met 79	1.00	Met 384	Val 52	0.22
Lys 386	Val 80	0.60	Met 384	Lys 55	0.26
Lys 386	Thr 82	0.30	Met 384	Val 56	0.65
Lys 386	Ala 83	0.39	Met 384	Thr 59	0.56
Thr 387	Thr 59	0.20	Met 384	Phe 76	0.29
Thr 387	Met 79	0.36	Met 384	Met 79	0.67
Glu 388	Gln 71	0.66	Phe 385	Leu 44	0.27
			Phe 385	Met 79	0.42
			Phe 385	Ala 83	0.48
			Lys 386	Met 79	0.28
			Thr 387	Thr 59	0.21
			Thr 387	Met 79	0.38

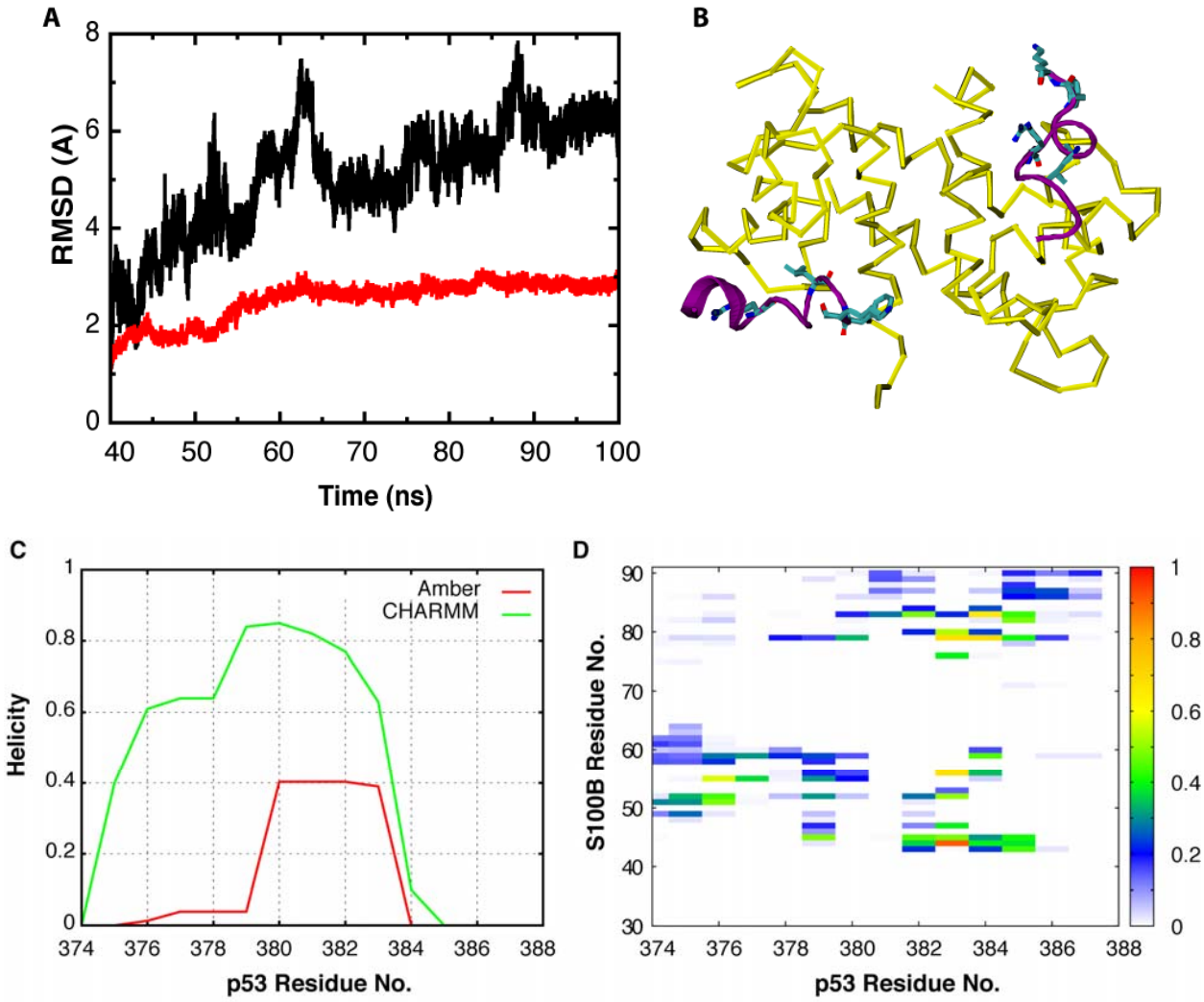


Figure S3. A) Cα RMSD of the whole complex (red) and first p53 monomer (black) from the initial conformation and B) the final snapshot from the 100-ns production simulation in the Amber99SB force field. Note that the RMSD values of p53 were calculated by first aligning the backbone heavy atoms of S100B(ββ) with the reference (initial) conformation (referred to as binding RMSD in the main text). C) Averaged residue helicity calculated by the Amber99SB trajectory (red trace) in comparison to the CHARMM ensemble (green trace, see Fig. 3 of the main text). D) Intermolecular residue contact maps computed from the Amber99SB trajectory. Both p53 monomers were included in calculation of the helicity profile and contact map shown.

Table S2: Summary of NOE violation analysis of the Amber99SB trajectory.

Atom Pairs	R_{nmr} (Å) [a]	R_{sim} (Å) [a]	Violation (Å)
L383 HD* - L3 HD* [b]	6.0	6.8	0.8
L383 HD* - M7 HE* [b]	5.0	5.8	0.8
Q375 HB* - E51 HG*	6.0	4.9	–
R379 HB* - K48 HG*	5.0	9.0	4.0
R379 HD* - V52 HG2*	6.0	3.3	–
L383 HD* - V52 HG2*	6.0	3.4	–
L383 HD* - V56 HG*	5.0	2.1	–
L383 HD* - V77 HG2*	6.0	6.8	0.8
L383 HD* - M79 HE*	3.3	2.7	–
L383 HB* - M79 HE*	5.0	4.0	–
L383 HD* - V80 HG2*	4.2	2.7	–
L383 HD* - V80 HG1*	5.0	2.8	–
L383 HD* - T82 HG2*	6.0	8.0	2.0
L383 HB* - A83 HB*	5.0	4.1	–
L383 HD* - A83 HB*	5.0	4.9	–
M384 HB* - K55 HG*	6.0	2.6	–
M384 HG* - V56 HG*	6.0	7.2	1.2
M384 HE* - V56 HG*	6.0	2.7	–
M384 HE* - L44 HD*	6.0	3.0	–
F385 HE* - M79 HE*	6.0	7.4	1.4
F385 HZ - M79 HE*	6.0	4.1	–
F385 HE* - V80 HG2*	6.0	5.1	–
F385 HZ - V80 HG2*	6.0	10.0	4.0
F385 HD* - V80 HG1*	6.0	7.1	1.1
F385 HE* - A83 HB*	5.0	8.0	2.0
F385 HZ - A83 HB*	6.0	4.4	–
F385 HZ - L44 HD*	6.0	6.1	1.1
K386 HD* - T82 HG2*	5.0	7.7	1.7
K386 HG* - T82 HG2*	6.0	4.8	–
K386 HD* - L44 HD*	6.0	5.0	–
T387 HG* - K55 HG*	5.0	7.3	2.3
E388 HG* - M79 HE*	5.0	8.9	3.9
E388 HG* - T82 HG2*	6.0	7.2	1.2

[a] R_{nmr}: experimentally assigned upper bound; R_{sim}: ensemble-averaged NOE-like distance calculated from the simulated ensemble.

[b] Involving p53-S100B(ββ) side chains across the symmetric dimer interface.

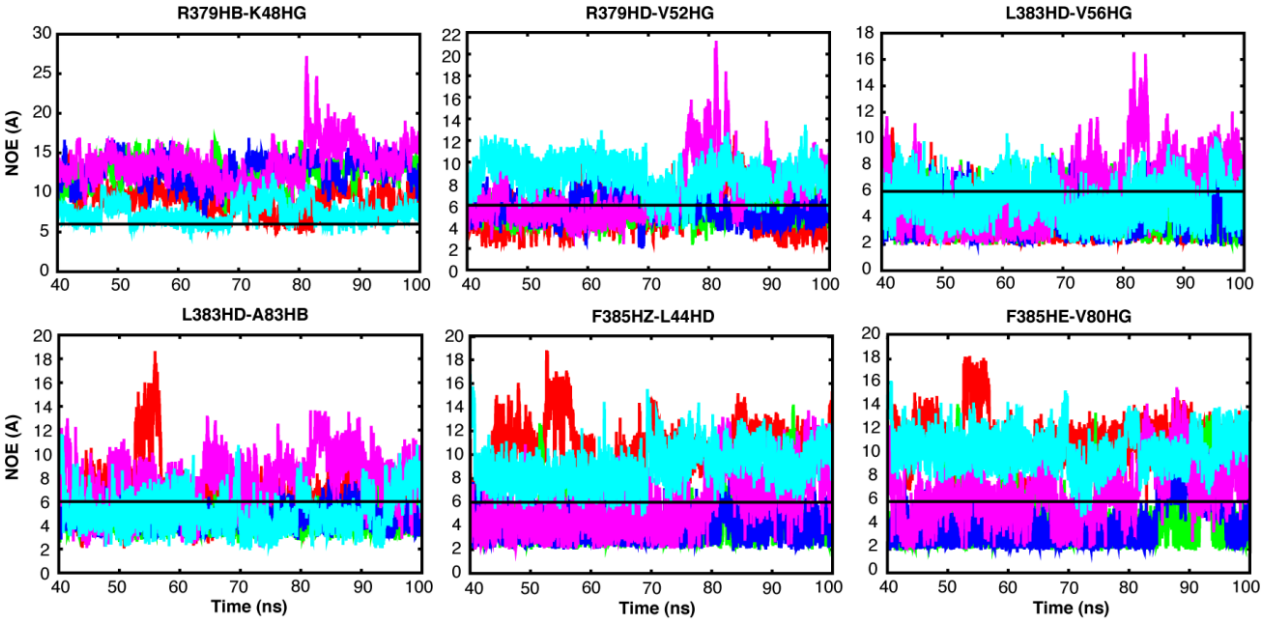


Figure S4. Time evolution of NOE-like distances between pairs of atom groups involved in six selected intermolecular NOEs during the last 60 ns of production simulations. The color-coding is the same as in Fig. 3A. The black lines mark the experimentally assigned NOE upper bounds.

Table S3: NOE distances calculated for monomer 1 from the last 60 ns of mode 1 trajectory using simple r^{-6} averaging (R_{sim}^{rigid}) and by considering the full effects of internal dynamics (see the following text for a brief summary of the analysis method). [a, b]: See Table S2.

Atom Pairs	R_{nmr} (Å) [a]	R_{sim}^{rigid} (Å)	S^2	τ_e (ps)	$R_{sim}^{dynamic} / R_{sim}^{rigid}$
L383 HD* - L3 HD* [b]	6.0	6.8	0.957	5516	1.006
L383 HD* - M7 HE* [b]	5.0	7.6	0.916	2783	1.014
Q375 HB* - E51 HG*	6.0	7.4	0.708	2792	1.055
R379 HB* - K48 HG*	5.0	6.4	0.706	3313	1.053
R379 HD* - V52 HG2*	6.0	2.9	0.521	1459	1.130
L383 HD* - V52 HG2*	6.0	3.0	0.580	3455	1.081
L383 HD* - V56 HG*	5.0	2.0	0.848	2176	1.028
L383 HD* - V77 HG2*	6.0	7.8	0.951	4216	1.007
L383 HD* - M79 HE*	3.3	2.8	0.643	1995	1.079
L383 HB* - M79 HE*	5.0	4.2	0.704	1979	1.062
L383 HD* - V80 HG2*	4.2	4.3	0.811	2789	1.033
L383 HD* - V80 HG1*	5.0	5.1	0.845	1272	1.032
L383 HD* - T82 HG2*	6.0	9.0	0.938	2426	1.010
L383 HB* - A83 HB*	5.0	5.7	0.876	3154	1.020
L383 HD* - A83 HB*	5.0	7.2	0.878	2599	1.021
M384 HB* - K55 HG*	6.0	7.0	0.798	2936	1.035
M384 HG* - V56 HG*	6.0	2.4	0.748	947.6	1.058
M384 HE* - V56 HG*	6.0	2.9	0.721	1647	1.060
M384 HE* - L44 HD*	6.0	4.3	0.823	3650	1.029
F385 HE* - M79 HE*	6.0	7.2	0.355	871.3	1.232
F385 HZ - M79 HE*	6.0	5.1	0.361	809.7	1.247
F385 HE* - V80 HG2*	6.0	5.2	0.432	974.0	1.189
F385 HZ - V80 HG2*	6.0	8.2	0.323	2171	1.209
F385 HD* - V80 HG1*	6.0	5.6	0.344	2310	1.191
F385 HE* - A83 HB*	5.0	4.0	0.412	1033	1.200
F385 HZ - A83 HB*	6.0	3.4	0.415	618.4	1.205
F385 HZ - L44 HD*	6.0	4.6	0.361	809.7	1.247
K386 HD* - T82 HG2*	5.0	3.3	0.832	1831	1.033
K386 HG* - T82 HG2*	6.0	3.4	0.782	2321	1.041
K386 HD* - L44 HD*	6.0	9.3	0.904	1467	1.018
T387 HG* - K55 HG*	5.0	9.0	0.815	2870	1.032
E388 HG* - M79 HE*	5.0	7.9	0.744	1025	1.059
E388 HG* - T82 HG2*	6.0	9.1	0.870	2213	1.023

Influence of internal dynamics on the calculated NOE distances

We largely followed the approach laid out in Bruschweiler *et al.* (1992) [1] for analyzing the influence of internal dynamics on the NOE distances calculated from the last 60 ns of the 100 ns CHARMM trajectory initiated from model 1 of the NMR ensemble (see the main text). The analysis was performed using a slightly modified version of the CHARMM NMR module [1,2]. Briefly, we assumed that the internal and overall motions of the protein are uncoupled, and only internal correlation functions corresponding to individual proton-proton vectors were calculated from the MD trajectory. Most intermolecular NOEs involve groups of multiple hydrogen atoms, particularly one or two methyl groups. Following the work of Schneider et al. (1999) [3], we represented each proton group using the heavy atom nearest to its center of mass to calculate a single internal correlation function for each pairs of proton groups. The generalized order parameters, S^2, and effective internal correlation time constants, τ_e, are given in Table S3. Given the spectral densities $J_{ij}(\omega)$, the cross-relaxation rates between spines *i* and *j* are given as,

$$\sigma_{ij}^{dynamic} = \frac{\pi}{5}\gamma^4 \mathsf{h}^2 \left[6J_{ij}(2\omega) - J_{ij}(0)\right], \qquad (1)$$

where ω is the Larmor frequency and γ is the gyromagnetic ratio of protons. In the limit of rigid proteins, the spectral density is given by the Fourier transform of the overall tumbling correlation function (assuming isotropic tumbling),

$$J_{ij}^{rigid}(\omega) = \frac{\left\langle r_{ij}^{-6} \right\rangle}{2\pi} \frac{\tau_R}{1+\omega^2\tau_R^2}\ , \qquad (2)$$

where τ_R is the overall molecular tumbling time constant. The corresponding cross-relaxation rates in the limit of rigid molecular, σ_{ij}^{rigid}, can be then calculated by substituting Eq. 2 into Eq. 1. The effects of internal dynamics on the back-calculated NOE distances are then estimated as:

$$\frac{R_{sim}^{dynamic}}{R_{sim}^{rigid}} = \left[\frac{\sigma_{ij}^{rigid}}{\sigma_{ij}^{dynamic}}\right]^{1/6}. \qquad (3)$$

The molecular tumbling time was set to be 12.0 ns based on the HYDRONMR analysis of the truncated PDB structure (model 1) [4]. The length of internal correlation functions was set to 12.0 ns (20% of the trajectory length). All relaxation rates were calculated assuming a 600 MHz proton frequency as used in the experimental study [5].

1. Brueschweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M., and Ernst, R. R., Influence of rapid intramolecular motion on NMR cross-relaxation rates. A molecular dynamics study of antamanide in solution. *J Am Chem Soc* **114** (7), 2289 (1992).

2. Chen, J. H., Brooks, C. L., and Wright, P. E., Model-free analysis of protein dynamics: assessment of accuracy and model selection protocols based on molecular dynamics simulation. *J Biomol NMR* **29** (3), 243 (2004).
3. Schneider, T. R., Brunger, A. T., and Nilges, M., Influence of internal dynamics on accuracy of protein NMR structures: derivation of realistic model distance data from a long molecular dynamics trajectory. *J Mol Biol* **285** (2), 727 (1999).
4. Garcia de la Torre, J., Huertas, M. L., and Carrasco, B., HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **147** (1), 138 (2000).
5. Rustandi, R. R., Baldisseri, D. M., and Weber, D. J., Structure of the negative regulatory domain of p53 bound to S100B(beta beta). *Nat Struct Biol* **7** (7), 570 (2000).