

# Dynamic Ligand-Induced-Fit Simulation via Enhanced Conformational Samplings and Ensemble Dockings: A Survivin Example

In-Hee Park<sup>†</sup> and Chenglong Li<sup>\*,†,‡</sup>

*Chemical Physics Program and Division of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, The Ohio State University, Columbus, Ohio 43210*

Received: November 21, 2009; Revised Manuscript Received: March 10, 2010

Survivin is an anticancer drug target due to its overexpression in tumor cells in a homodimer form. Abbott Laboratories has identified a small molecule binding site near the dimerization interface in a high-throughput-screening (HTS)-NMR experiment. A benchmarking of the binding mode of the compound Abbott8 aided in the search for the ligand-induced-fit receptor structure by exploring the conformational space of the survivin dimer. We performed ensemble dockings with Abbott8 against a large set of conformations sampled via replica exchange molecular dynamics (REMD). This enhanced sampling allowed the reproduction of the holo-NMR experimental binding mode. Surprisingly, the major structural change in the best-REMD snapshot corresponding to the small molecule induced-fit happens in the so-called “survivin mitosis/apoptosis switch loop”, consistent with the X-ray crystal structure of survivin-monomer/borealin/INCENP chromosomal passenger complex (CPC), as the distance between Phe93 and Phe101 increased. To verify this hypothetical pathway for the induced-fit conformational change, we utilized morphed intermediate structures that combined the X-ray data and the best-REMD snapshot, and the potential of mean force (PMF) of the survivin dimer was constructed with umbrella sampling (US) followed by a multiple Bennett acceptance ratio estimator (MBAR). It revealed a 3–4 kcal/mol free energy barrier along the reaction coordinate, and the complex is stabilized by the gain of the binding energies of Abbott8. This free energy barrier might prohibit the reproduction of the experimental binding mode from the regular NTP-MD ensemble docking that we had tried. The combination of REMD generalized ensemble sampling with ensemble docking and free energy pathway analysis may provide a novel research protocol for the simulation of protein–ligand induced-fit recognition.

## Introduction

Survivin is a member of the inhibitor of apoptosis protein (IAP) family. Its dimerization typically occurs as two homogeneous monomers that are symmetrically held together by hydrophobic residues to form a bow-tie shaped dimer, which suppresses programmed-cell-death (i.e., apoptosis) and can result in the proliferation of tumor cells.<sup>1,2</sup> In contrast, the monomeric form of survivin plays a role in normal cell division (i.e., mitosis) as part of the chromosomal passenger complex by packing with Borealin and INCENP along their long C-terminal  $\alpha$  helix structures.<sup>3,4</sup>

The survivin homodimer has two distinct binding hotspots, the conventional Baculovirus IAP Repeat (BIR) domain Smac N-term binding cleft and the recently identified dimerization interface (Figure 1A).<sup>5</sup> We will probe the conformational “plasticity” in the dimerization interface with the aim of designing and discovering dimerization inhibitors for anticancer applications. Therapeutic approaches to minimizing the anti-apoptotic function of the survivin dimer have traditionally focused on the Smac-peptide binding site in the BIR-domain.<sup>6,7</sup> However, in 2007, Abbott Laboratories identified a potential small-molecule binding site near the dimerization interface that is distinct from the Smac-peptide binding site.<sup>5</sup> They also used two types of affinity-based screening methods, HTS-NMR and AS/MS, to discover a new set of small-molecule inhibitors that

possibly disrupts survivin dimerization. Since then, the direct disruption of dimerization via small-molecule binding has become a potential strategy.<sup>8</sup>

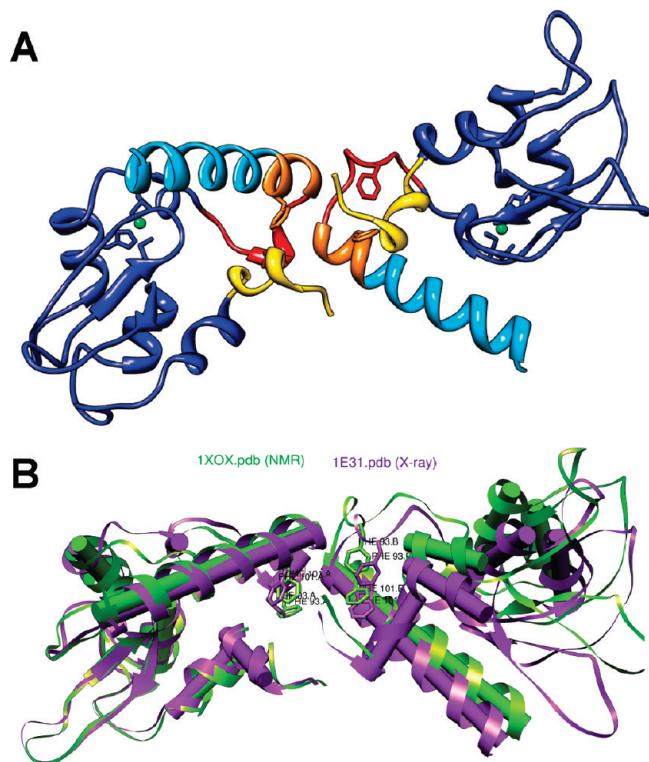
Upon the discovery of a set of survivin interface binders, a characterization of compound Abbott8’s binding modes aided in the refinement of the receptor conformation by defining the conformational flexibility of survivin. In our initial effort, we had tried both flexible docking of Abbott8 to the X-ray structure (PDB code 1E31)<sup>1</sup> of survivin with key residues (Phe93 and Phe101) flexible and ensemble dockings to fit the Abbott8 compound to the survivin dimer snapshots collected from the NTP-MD simulation. However, the reproduction of the holo-NMR (Abbott8 compound + survivin dimer complex) binding mode failed. The resolved X-ray structure differs from the apo-NMR (PDB code 1XOX)<sup>9</sup> structures at the interface (Figure 1B). The ensuing docking of Abbott8 to apo-NMR structures was not successful, too. Therefore, we used the enhanced sampling method of replica exchange MD (REMD). This allowed us to identify the optimal conformation of survivin suitable for ligand binding, in which the Abbott8 compound was used as a probe to dock against a large set of REMD snapshots.

To further gain the insights on the ligand-induced-fit recognition mechanism during the binding process, we utilized morphed intermediate structures spanning from X-ray structure to the reproduced holo-conformation snapshot from the REMD. We then identified the key reaction coordinates by comparing the distance matrix variations. To verify the chosen reaction coordinate, we calculated the binding energy of Abbott8 to the

\* To whom correspondence should be addressed. Phone: (614) 247-8786. Fax: (614) 292-2435. E-mail: li.728@osu.edu.

<sup>†</sup> Chemical Physics Program.

<sup>‡</sup> Division of Medicinal Chemistry and Pharmacognosy.



**Figure 1.** (A) The homodimer structure of survivin ( $\Delta 5-117$ ), color-coded by region. It consists of a BIR-domain ( $\Delta 15-89$ , shown in blue) with the zinc-binding motif shown in ball-and-stick representation; a dimerization interface region (N-terminal  $\Delta 6-14$  portion, shown in yellow); a mitosis/apoptosis-switch loop ( $\Delta 89-97$ , shown in red) and the remaining interface ( $\Delta 98-102$  shown in orange); two key residues, Phe93 and Phe101, are shown in stick representation, and an amphiphilic C-terminal helix ( $\Delta 103-117$ , shown in sky blue). (B) Superimposed structures of the X-ray and NMR resolved survivin dimers with respect to one monomer side.

morphed intermediate structures via MMGBSA (molecular mechanics generalized born surface area). To reconfirm the hypothetical conformational recognition pathway, we reconstructed the potential of mean force (PMF) of the survivin dimer, cross-validated the reaction coordinate span again via umbrella sampling, and extracted the free energy path via both multistate Bennett acceptance ratio (MBAR) and weighted histogram analysis method (WHAM). The PMF profile shows an energy barrier that explains the necessity of this enhanced sampling method and the underlying physics of induced-fit mechanism.

## Materials and Methods

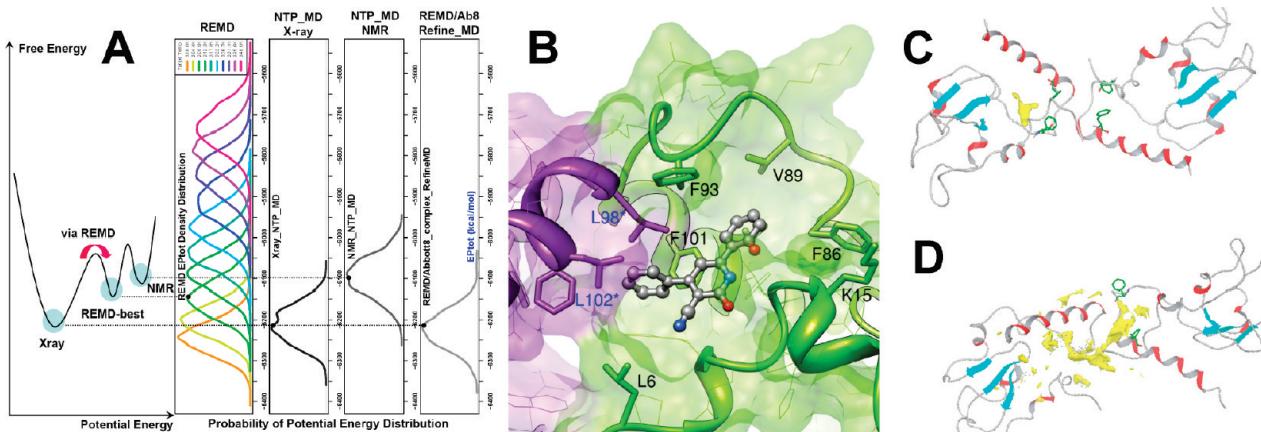
**Retrieval of Survivin Dimer Structure.** Residues  $\Delta 5-117$  of each survivin monomer were taken from the X-ray structure (PDB code 1E31) for the regular NTP-MD and REMD simulations<sup>1</sup> and from the apo-NMR structure (PDB code 1XOX) for the regular NTP-MD.<sup>9</sup> Residues  $\Delta 118-143$  from the long amphiphilic C-terminal  $\alpha$  helix of each monomer were truncated to reduce the size of the solvated water-shell. This C-terminal truncation strategy was also used in experimental studies to increase its solubility but did not affect dimerization.<sup>9</sup> In addition, to maintain the overall stability of the survivin, the tetrahedral zinc binding structure of each BIR domain was coordinated with two deprotonated Cys (residue type CYM in the Amber convention)<sup>10</sup> and one His. The position of the His protonation site was selected to avoid disruption of the unpaired electron coordinates around the zinc atom: the residue type HID (protonated at  $\delta$ -position nitrogen atom) was used for the X-ray

structure; the residue type HIE (protonated at  $\epsilon$ -position nitrogen atom) was used for the apo-NMR structure.

**REMD Setup for the Survivin Dimer.** We are interested in a small scale conformational change at the dimerization region, which requires different parameters than do conventional folding/unfolding studies. To meet this need, we chose lower and smaller temperature ranges (300–340 K) and set up the REMD system as follows. *Step 1:* the survivin dimer system was solvated with a 15 Å thick truncated octahedron TIP3P water model, neutralized by  $\text{Na}^+$  ions, with minimized energy. The energy minimized structure was used for all replicas in the explicit water solvated equilibration stage. *Step 2:* the number of replicas ( $M$ ) needed to cover the temperature range of 300–340 K was estimated by  $M \approx \sqrt{f} \times \ln(T_{\max}/T_{\min})$ , where  $f$  is the total degrees of freedom associated with the 228 total residues of the survivin dimer system.<sup>11,12</sup> The final estimation of  $M$  was approximately 6, but we set up total 10 replicas at 4 K temperature intervals to ensure efficient replica exchange rate to fulfill the convergence criteria. *Step 3:* explicit water solvated systems for 10 replicas were carried out for 50 ps temperature equilibration followed by 200 ps pressure equilibration at ten different target temperatures using Berendsen coupling schemes in Amber9. *Step 4:* The equilibrated explicit water solvated systems were converted to implicit water systems with  $\text{GB}^{\text{OBC}}$  by manipulating the restart file (containing the coordinate and velocity information of all protein and water atoms) of the latest snapshot of the equilibration stage from each replica. We extracted the coordinate and velocity information of only the survivin dimer atoms. Finally, an implicit water topology file was created for the REMD run. *Step 5:* in the main temperature exchange production run, each replica was set to attempt to exchange its temperature with other replica 10 000 times at every 0.2 ps, resulting in 20 ns trajectories from the ten replicas. REMD was performed on the IBM Opteron cluster using 40 cores (4 processors per replica).

**Ensemble Dockings.** The Abbott8 compound was used as a probe ligand and is expected to recognize the ligand-induced-fit conformation of the survivin dimerization interface. For the fast dockings to the large set of survivin dimer snapshots, we used Glide in the Schrödinger package.<sup>13</sup> *Step 1:* the REMD trajectory was extracted as a PDB file. *Step 2:* every PDB was converted to the maestro format by PrepWizard. *Step 3:* a docking grid file covering the dimerization interface site was generated for each snapshot. *Step 4:* a series of Glide dockings with XP (extra precision) scoring function to all prepared grid files of each snapshot was carried out. *Step 5:* the resulting docking modes were analyzed based on the distance between the F93 side chain and phenol ring of Abbott8 compound.

**Morphed Structure-Based Ensemble Dockings and Binding Energy Calculation.** To identify the reaction coordinate, residue-wise distances were measured for the selected REMD snapshot, the X-ray structure, and the apo-NMR structure via the *matrix dist* protocol of AmberTools1.2.<sup>10</sup> The distance matrix subtractions and resulting contour plots were subsequently analyzed with the R statistical computing package.<sup>14</sup> The X-ray structure of the survivin dimer was superimposed on the best-reproducible REMD snapshot with respect to the monomer-A side. The overall structural difference was most evident in the F93–F101 distance at the monomer-B side, including the N-terminal position. Then a total of nine morphed intermediate structures were generated starting from the X-ray as a reference structure and transitioning toward the best-REMD snapshot via Chimera,<sup>15</sup> each of which was minimized using the Amber force field. The Abbott8 compound was docked to each of the



**Figure 2.** (A) Schematic potential energy profile of the survivin dimer with corresponding MMGB (molecular mechanics part + implicit Generalized Born-model based solvation energy) potential energy distributions collected from the REMD (potential energy range of  $-6400$  to  $-5500$  kcal/mol from ten target temperatures), NTP-MD of X-ray, NTP-MD of NMR, and refined MD of the best-REMD/Abbott8 complex. (B) A reproducible binding mode of Abbott8 at the survivin dimerization interface. This receptor conformation (best-REMD snapshot) was found in the REMD sampling batch at the target temperature of  $308.9$  K. (C) Hydrophobic patch distribution on the candidate binding site evaluated for the X-ray structure by SiteMap. (D) The best-REMD snapshot.

morphed intermediate structures, including reference and target structures, via Glide. The docked coordinate of the Abbott8 compound was concatenated with the corresponding morphed survivin structure in amber trajectory format. Binding energies were recomputed via MMGBSA protocol.

**Morphed Structure-Based Umbrella Sampling and Reconstruction of PMF via MBAR and WHAM.** The F93–F101 distance was selected as a reaction coordinate and a total of 24 umbrella samplings were carried out in Amber9. The morphed intermediate structures were used as initial coordinates for the umbrella sampling MDs. The reaction coordinate was varied from  $4.5$  to  $12.5$  Å by  $0.3$ – $0.4$  Å window intervals with a harmonic constraint ( $k$ ) of  $40$  kcal/mol/Å $^2$ . The force constant was finally chosen based on the inspection of the probability distribution overlap among 24 umbrella sampling results. Reaction coordinate overlap was checked during restrained explicit water US-MDs in the minimization and equilibration stages, as well as during a 1 ns production run, by writing out the distance information at every 20 fs. Finally, the PMF was constructed from the reaction coordinate sampled via multiple-US MDs with two free energy estimators (MBAR<sup>16</sup> and WHAM<sup>17</sup>) using the appropriate force constant units.

## Results

**Reproduction of the Holo-NMR Binding Mode of Abbott8 and Experimental Binding Energy. Assessment of REMD Sampling Space.** The conformational sampling space for the survivin dimer was enhanced via REMD, which was probed in terms of potential energy distribution ranges against its counterparts via the NTP-MDs shown in Figure 2A. Sufficient overlap between the adjacent replicas may ensure sampling convergence. Explicit-water solvated equilibration followed by an implicit-solvent REMD production run for the survivin dimer gave an acceptance ratio of 30–40% for the ten replicas.

We then identified the most reproducible binding mode of Abbott8 to the holo-NMR experiment in the ensemble docking results and found that its complementary receptor structure (called best-REMD snapshot hereafter) belonged to the REMD pseudotrajectory at a target temperature of  $308.9$  K. For survivin alone, relative potential energies can be deduced based on the ensemble distributions of the  $308.9$  K REMD and the NTP-MDs of X-ray and apo-NMR: the X-ray structure has the most

stable conformation, while the best-REMD conformation is less stable than the X-ray structure but more stable than the apo-NMR structure schematically depicted in Figure 2A. For the survivin/Abbott8 complex, the potential energy distribution from the NTP-MD of the best-REMD/Abbott8 structure was lowered near the X-ray ensemble distribution, which indicated that the Abbott8 compound did indeed induce a more stabilized structure.

**Intermolecular Interactions between Abbott8 and Residues at the Dimerization Interface.** The phenol ring of the Abbott8 compound penetrated the binding pocket on the monomer-B side. It is aligned nearly in parallel to the aromatic ring side chain of Phe93. In addition, the benzene ring of the Abbott8 compound touched the pocket on monomer-A (the other monomer in the dimer) formed by Leu98 (monomer-A) and Leu6 (monomer-B) as shown in Figure 2B. The relative side chain arrangement around the Abbott8 compound was distinct from that in the X-ray structure but similar to that in the holo-NMR structure. Here we noticed that the apo-NMR structure of the survivin dimer was obtained by resolving the single monomer structure, followed by protein–protein docking with NOE signal distance restraints.<sup>9</sup> We observed that the docked dimer structure of apo-NMR lost its tertiary structure over a 5 ns NTP-MD at  $300$  K. In contrast, the X-ray structure retained its tertiary structure under these conditions. We used the X-ray structure for the REMD because it was fully resolved for both monomers, and thus the spatial arrangement between monomers was more reliable in the X-ray structure than in the apo-NMR structure. Therefore, it is meaningful physically that the ligand-induced conformation was generated starting from fully resolved X-ray dimer structure through REMD and identified via ensemble dockings.

**Binding Stability and Ensemble Binding Energy via MD Refinement.** Several conformations from the REMD followed by ensemble docking showed a binding mode similar to that in the holo-NMR structure. To ensure the stability of binding in the binding pocket, we performed a short MD to discriminate stable binding modes from the single, static docking modes that immediately lost their initial docking coordinate early in the MD steps. The final candidate binding mode was simulated up to 5 ns NTP-MD with its corresponding receptor structure, and the ensemble binding free energy was computed via MMGBSA protocol. The calculated ensemble binding energy of the Abbott8

**TABLE 1: Calculated Binding Energy of Abbott8 to the Survivin Dimer Interface Averaged over 5 ns: MMGBSA versus Experimental Binding Energy**

energy component	binding energy (kcal/mol)
$\Delta E^{\text{ele}}$	$-6.79 \pm 4.01$
$\Delta E^{\text{vdw}}$	$-35.85 \pm 3.05$
$\Delta G_{\text{sol}}^{\text{ele}}$	$21.80 \pm 3.93$
$\Delta G_{\text{sol}}^{\text{np}}$	$-5.20 \pm 0.24$
$\Delta G_{\text{GBSA}}^{\text{MM}}$	$-26.03 \pm 2.58$
$T\Delta g^{\text{tot}}$	$-20.48 \pm 15.10$
$\Delta G^{\text{calc}}$	$-5.55 \pm 15.32$
$\Delta G^{\text{exp}}$	$-5.58 \pm \text{N/A}$

compound was  $-5.6$  kcal/mol, which is also similar to the experimental binding energy (Table 1). The van der Waals energy component contributed most of the final binding energy, which may be associated with the ring stacking binding mode in the hydrophobic pocket.

**Binding Hotspot Distribution upon Conformational Change.** The candidate binding site for the best-REMD snapshot was identified via SiteMap, which searches for sites suitable for ligand binding, produces hydrophobic and hydrophilic maps, and then evaluates the probable binding site by assessing properties such as extent of exposure to the solvent, hydrophilic/hydrophobic character, and pocket volume.<sup>18</sup> The hydrophobic site at the dimerization interface region was increased in the best-REMD snapshot relative to the X-ray structure (Figure 2C,D). The optimal REMD snapshot we found is indeed suitable for the hydrophobic ligand binding.

**Reaction Coordinate Definition and Construction of PMF from the REMD Trajectory.** The best-REMD snapshot found through REMD/ensemble docking preserved the BIR domain structure. The most significant structural variations were observed at the dimerization interface. We can therefore focus on the dimerization interface variation of the best-REMD with respect to the X-ray and apo-NMR structures to identify the reaction coordinate that captures the critical structural change.

**Distance Matrix Variation to Identify the Reaction Coordinate.** The dimerization interface can be separated into a mitosis/apoptosis switch loop ( $\Delta 89-102$ ) and N-terminal portion ( $\Delta 6-14$ ), as depicted in Figure 1A. To extract the key variation between the best-REMD snapshot and the reference structures of X-ray and apo-NMR, we first measured the residue–residue distances for the interface region and subtracted their distances from the other structure. This revealed that the most altered distance was Phe93, which moved toward the Leu98–Leu102 in the best-REMD snapshot and moved away from the original contact observed in the X-ray structure (Figure 3A). Wendt and colleagues also noted that the side chain rearrangement of Phe93 was the key difference between the holo-NMR and apo-NMR structures.<sup>5</sup> In addition, the N-terminal position was altered in the best-REMD structure with respect to the X-ray and NMR structures. Leu6 approached the mitosis/apoptosis-switch loop region at Gln92, while Trp10 moved away from the switch loop (Figure 3B). As shown in the third panels in Figure 3A,B, the distance variation in the interface region between X-ray and NMR was not significant. Although the optimal conformation is a result of interplay between F93–F101 distance and N-terminal conformation, the N-terminus is a loop structure and is thus intrinsically more flexible than the switch-loop and is not well-resolved in either the NMR or X-ray structure. Moreover, the key intermolecular interaction involving aromatic stacking occurred mainly at the switch-loop, which directly interacts with the Abbott8 compound. Therefore, we

defined the F93–F101 distance at the switch-loop as the critical reaction coordinate for the rest of this work.

**Conformational Change Sequence in Terms of PMF Energy Distribution.** Although the two survivin monomers are homogeneous, subtle conformational differences in the original X-ray structure suggest that they may behave differently. To investigate this difference, we used the same reaction coordinate of F93–F101 distance for both monomers and then constructed a 2D-PMF by analyzing the pseudotrajectory of REMD at target temperature 308.9 K (where the best-REMD snapshot was found) via WHAM and MBAR (Figure 4). The 2D-PMF results suggest a possible sequence for opening the dimerization interface pocket, with one monomer opening first while the other monomer maintains contacts similar to those seen in the X-ray structure. The diagonal path has relatively high PMF, which implied that opening both sides simultaneously is more difficult than opening one side at a time. The monomer-B side can be opened gradually while the monomer-A side more or less preserves the original contact as designated by two arrows.

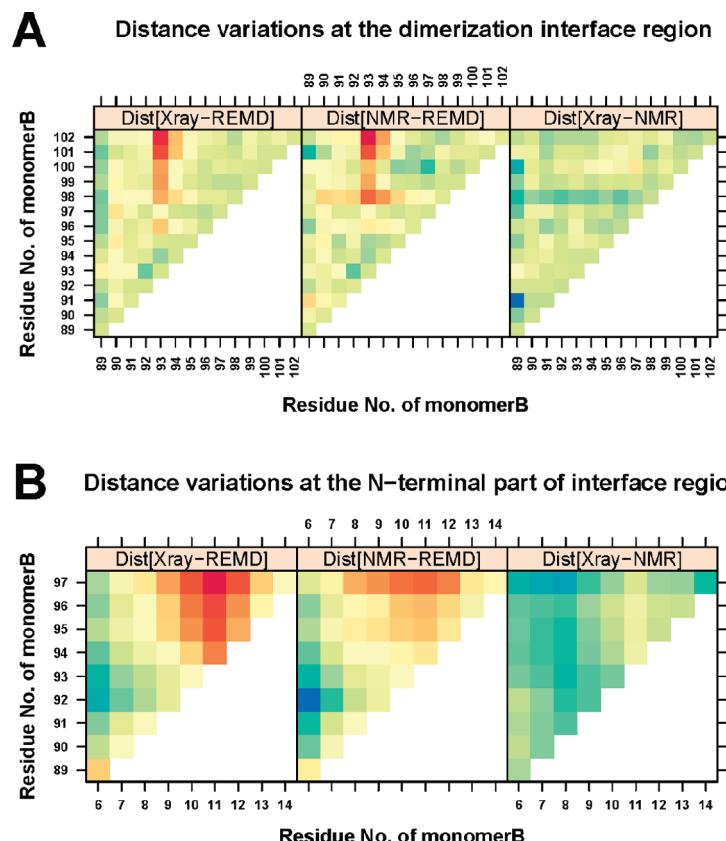
The final confirmed best-REMD snapshot has F93–F101 distances of 5.9 and 12.5 Å on the monomer-A and monomer-B sides, respectively. The region marked by a blue rectangular box denotes the selected reaction coordinate of F93–F101 distance of monomer-B side.

We determined the reaction coordinate via distance subtraction on the interface region (Figure 3), together with the relative conformational potential energy shown in Figure 4. We can then predict the conformational change pathway from the stable X-ray structure (4.5 Å; closed contact) toward the best-REMD snapshot (12.5 Å; opened) along the defined reaction coordinate of F93–F101 distance on the monomer-B side.

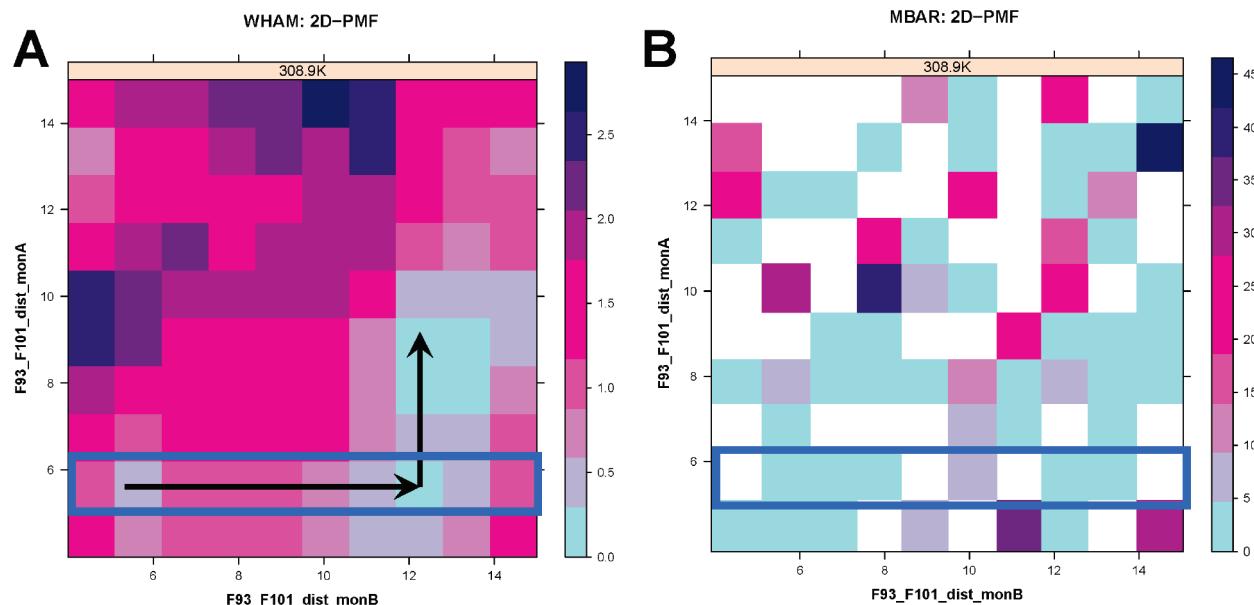
**Morphed Structure-Based Binding Energy of Abbott8 along the Reaction Coordinate.** To investigate the ligand-induced conformational change at the survivin dimerization interface, we utilized the hypothetical pathway along the chosen reaction coordinate. We first generated morphed intermediate structures connecting the energetically stable X-ray structure to the best-REMD snapshot in silico. Then to verify the hypothetical pathway by providing a physicochemical basis, we docked the probe molecule (the Abbott8 compound) to the morphed intermediate structures and calculated the binding energy of Abbott8 at each point.

**Opening of the Hydrophobic Channel along the Reaction Coordinate.** As the F93–F101 distance is increased, the hydrophobic surface was exposed so that the binding site became suitable to accept the Abbott8 compound (Figure 5A). A surface-capping view captured the penetration of the Abbott8 into the hydrophobic binding site, which was blocked in the X-ray structure, and then gradually opened toward the best-REMD snapshot (Figure 5B).

**Driving Forces and Intermolecular Interactions for Abbott8 Recognition along the Reaction Coordinate.** MMGBSA free energy analysis reveals the Abbott8-survivin induced-fit pathway. Initially, Abbott8 compound stuck on the surface due to the steric blockage; however, electrostatic binding force steers the dynamic binding process. The energy valleys in the electrostatic energy component (see ELE plot at Figure 6) showed that a specific hydrogen bond was formed between Abbott8 compound and the survivin dimer. This interaction induced the Abbott8 compound toward the right track for the dimerization interface. The phenol group of the Abbott8 compound formed a hydrogen bond with Arg18 near the N-terminal region, then formed a hydrogen bond with Glu40 at 6.3 Å F93–F101 distance, and then formed two hydrogen



**Figure 3.** Distance matrix subtraction representations for the structural variations at the dimerization interface of monomer-B. Structural variations of the best-REMD snapshot are given with respect to the X-ray and the NMR structures as well as the intrinsic difference between X-ray and NMR. Variations at the mitosis/apoptosis-switch loop region  $\Delta 89-102$  (A) and N-terminal interface region  $\Delta 6-14$  (B) are given. Only the lower triangle portion of the matrix is presented.

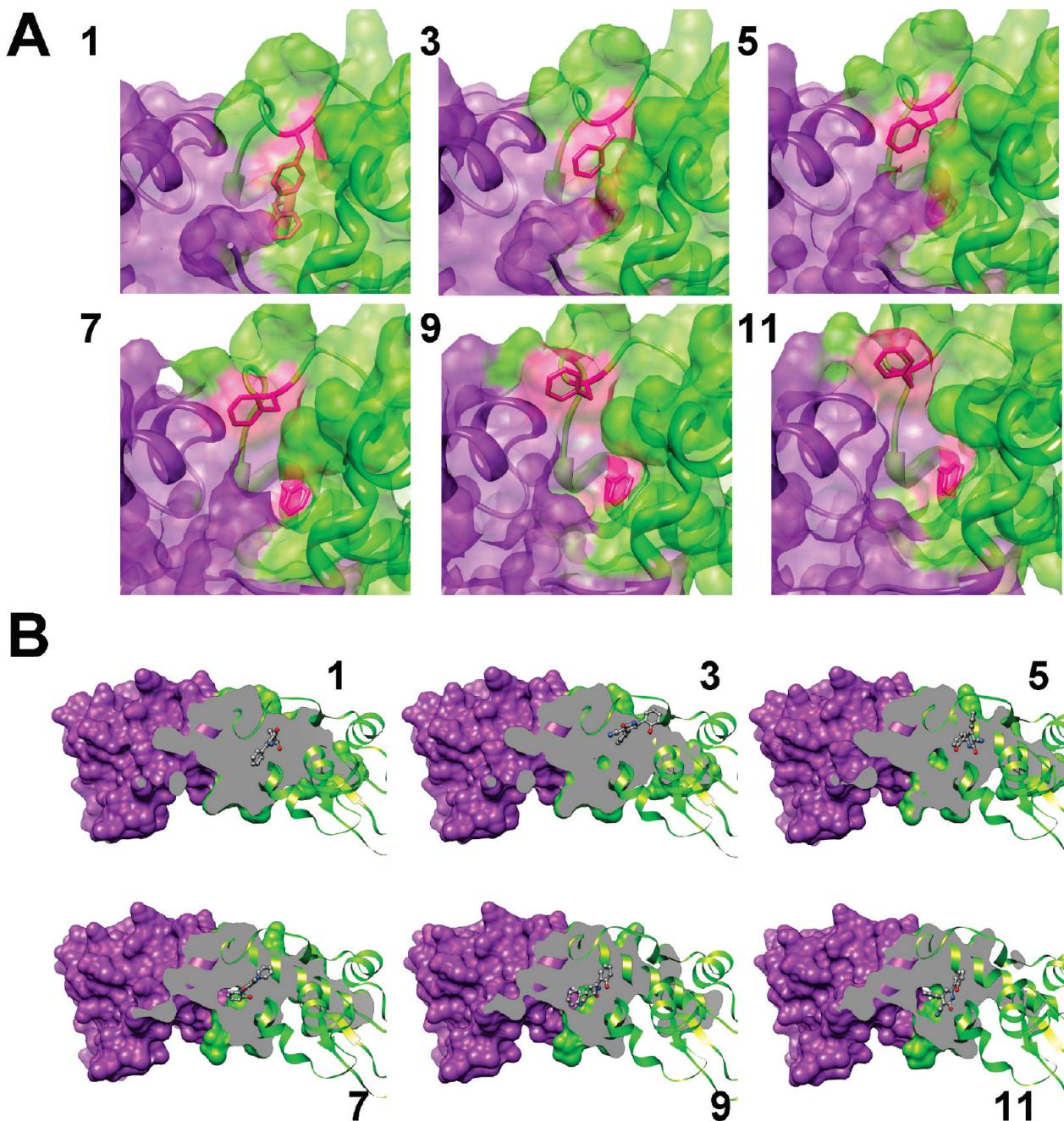


**Figure 4.** The 2D-PMF with distances between F93–F101 on monomer-A and monomer-B at 308.9 K estimated via WHAM (A) and via MBAR (B). The PMF is given in kcal/mol with a color bar legend on the right. The region marked by a blue rectangular box denotes the critical reaction coordinate. The empty (white) cells on the 2D-PMF via MBAR indicate that the corresponding bin size is zero, that is, no corresponding reaction coordinate was sampled.

bonds near the N-terminus between the phenol group and Phe13, and between nitrogen atom and Arg18 at 10.4 Å. From the edge of BIR domain toward the N-terminal interface region, electrostatic forces guided the Abbott8 compound to the hypothetical pathway by forming hydrogen bonds between the surrounding residues and hydrophilic groups of Abbott8. Once Abbott8

approached the interface, the van der Waals forces (see VDW panel in Figure 6) between the aromatic ring of Abbott8 and hydrophobic residues such as Phe93 and Phe101 played a role in the further stabilization of binding.

The hypothetical pathway was verified as a feasible route via continuous stabilization of the binding energy of Abbott8



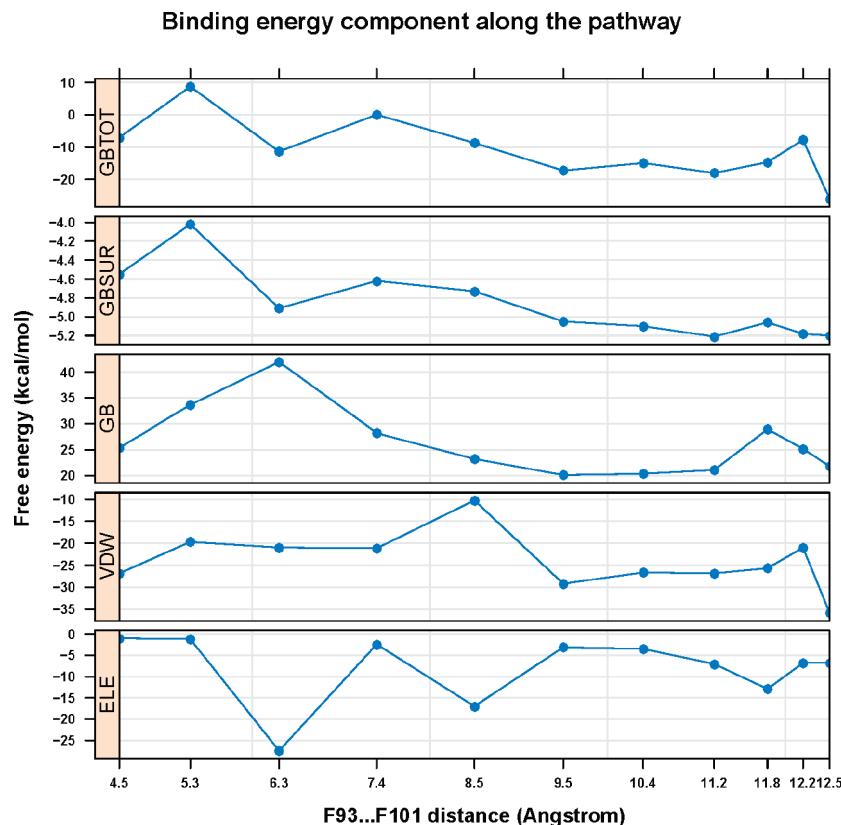
**Figure 5.** (A) A hypothetical conformational change pathway beginning from the X-ray structure (reference; snapshot number 1) through the morphed intermediate structures (3, 5, 7, and 9) and toward the best-REMD structure (target; snapshot number 11). The key Phe93 and Phe101 residues of the monomer-B side are represented in deep pink. (B) A surface cap representation showed that the binding site was closed in X-ray structure, started to open gradually in intermediate structures, and is fully open and suitable for Abbott8 binding in the best-REMD snapshot.

along the reaction coordinate. If we construct the PMF of the survivin dimer along the same reaction coordinate, we may expect to see an energy barrier along this pathway that may prevent the appropriate sampling via regular NTP-MD.

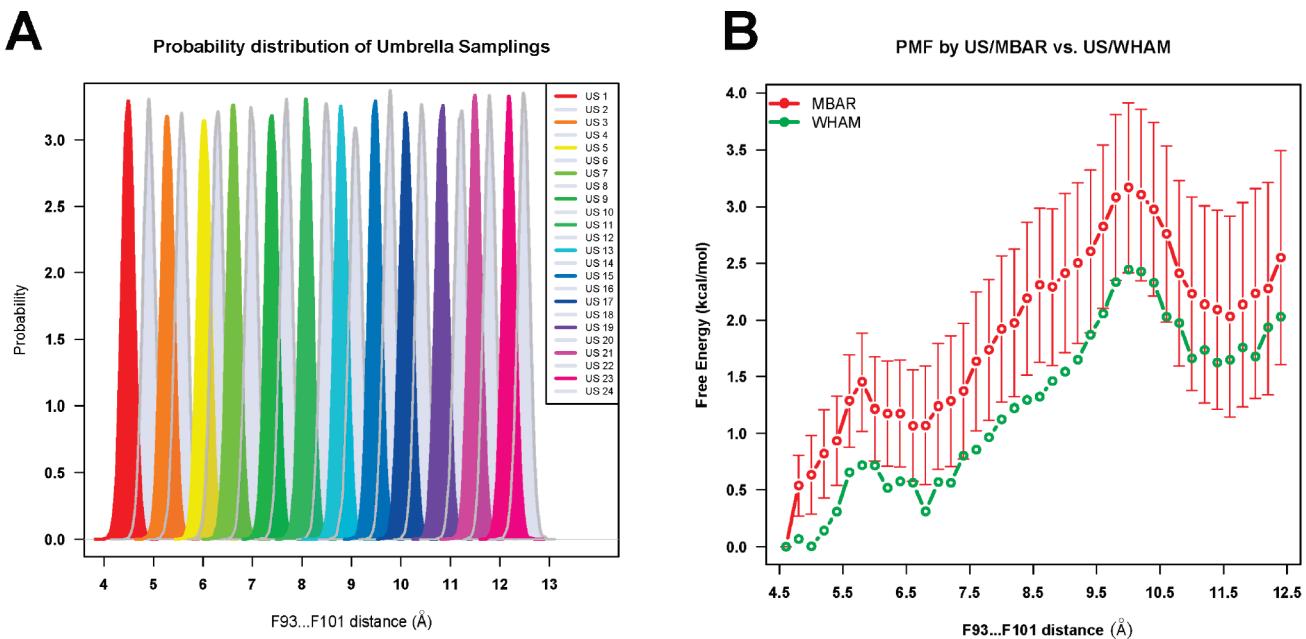
**Reconstruction of the PMF of the Survivin Dimer along the Reaction Coordinate.** To cross-validate the hypothetical pathway investigated by REMD and morphed structure-based binding energy of Abbott8, we reconstructed the PMF of the survivin dimer along the same reaction coordinate via umbrella sampling (US) with 24 windows using a 40 kcal/mol/Å<sup>2</sup> harmonic force constant. On the basis of the sufficient overall sampling shown in Figure 7A, we constructed the PMF via two different methods of removing the biased potential and extract free energy information from the multiple-equilibrium MDs

(Figure 7B). MBAR can also provide the standard deviation of each individual bin. Overall, both MBAR and WHAM showed similar free energy profiles; an approximately 3–4 kcal/mol free energy barrier occurred at 10 Å F93–F101 distance, which was consistent with the PMF results of REMD in Figure 4.

**Normal Mode Analysis (NMA) to Access Essential Dynamics of Survivin upon Abbott8 Binding.** To clarify the relationship between a ligand-induced large conformational change and a selected distance between two Phe residues as a reaction coordinate, we have additionally carried out a principal component analysis (PCA; also called essential dynamics<sup>22</sup>). We have combined the two separate MD trajectories (one for the apo-survivin MD, the other for the holo-survivin MD) and then calculated the eigenvectors via covariance matrices from the



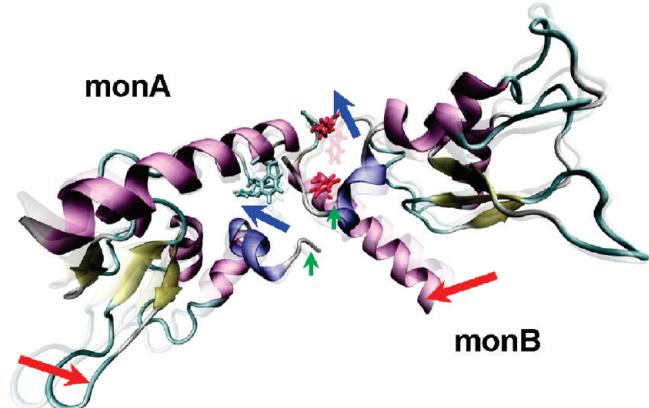
**Figure 6.** MMGBSA binding free energies of Abbott8 compound docked to the X-ray structure, nine morphed intermediate structures, and the best-REMD snapshot along the F93–F101 distance opening pathway. The final best-REMD snapshot energy was incorporated from the ensemble energy results from Table 1. Total binding energy (GBTOT) and deconvoluted energy components of nonpolar salvation energy, polar solvation energy, van der Waals, and electrostatic energy (GBSUR, GB, VDW, and ELE, respectively) are shown.



**Figure 7.** (A) Umbrella sampling probability distributions from 24 windows. (B) PMF along the reaction coordinate (F93–F101 distance on monomer-B) evaluated by MBAR and WHAM.

combined trajectories (see the Supporting Information for detailed method). The first three eigenvectors possessing a high percentage of the significant molecular motion are selected. Then, the projections of the trajectories onto these selected eigenvectors are viewed as a deviation from the averaged structure to the directional movements, as depicted in Figure 8. The three distinct ensemble dynamic modes are (1) Phe93–Phe101

distance increment from the closed form to the opened form, denoted by blue arrows; (2) N-termini ordered arrangement toward the dimerization interface, denoted by green arrows; and (3) a relative adjustment of the two monomers' coordinates, marked by red arrows, which indicates a global adjustment. The single distance opening between two Phe residues may occur only with an extensive conformational change, such as domain



**Figure 8.** Essential dynamics projected onto the first three eigenvectors based on combined MD trajectories of apo-survivin MD and holosurvivin MD. Three set of arrows depict the deviation of principal components from the averaged structure (faded color).

adjustment. Therefore, although we have selected the Phe93–Phe101 distance as a reaction coordinate, it is presumably a result of a collective large-scale movement in the course of dimerization interface adjustment. Also, it is noteworthy that the large-scale domain adjustment of the two monomers is similar to the already observed intrinsic structural difference between the X-ray and NMR structures of the survivin dimer, shown in Figure 1B.

## Discussion

**Explicit Water-Solvated Equilibration Followed by Continuum Solvent REMD.** Given the small contact area between survivin monomers, which are held together by hydrophobic residues at the dimerization interface, the resulting acceptance ratio of our REMD was fairly high (approximately 30–40% for all replicas). Other than the small temperature space between adjacent replicas (4 K), an explicit water solvated equilibration strategy used prior to the continuum solvent REMD production run improved the acceptance ratio. In addition, during the entire temperature and pressure equilibration period, the two monomers kept their hydrophobic contact close at the dimerization interface without an aid of distance restraint, even at maximum temperature of 340 K. Most of all, an explicit water solvated equilibration prevented the artificial structural denaturation such as overstabilization of ion pairs and secondary structure bias<sup>19</sup> that may be caused by inaccuracies in the continuum solvent model. In contrast, the explicit water model explicitly balances protein–protein and protein–water interactions in the equilibration stage. To improve the convergence and remove the secondary structure bias caused by the continuum solvent (e.g., GB) model, and at the same time to reduce the degrees of freedom in the water molecules, a hybrid model was developed by considering the first layer of the water shell. This attempt indicated that the explicit water model is needed for better sampling.<sup>19</sup> For a system as large as the survivin dimer, hybrid REMD is intractable due to huge number of degrees of freedom even from the first layer water molecules. Therefore, an explicit water equilibration followed by GB-REMD may be an alternative to hybrid REMD for large macromolecular systems.

**REMD As a Tool for Structure Refinement.** We have used a narrow range of low temperatures to obtain small-scale conformational changes, which is in contrast to the conventional application of REMD for observing protein folding/unfolding. With the aid of an optimal conformational recognition procedure via ensemble dockings against a large set of multiple receptor

conformations through REMD, we identified the best induced-fit conformation. In drug design, either high-resolution NMR or X-ray structures have been thought to be prerequisite to initiation of virtual screening. In survivin dimer case, although both NMR and X-ray structures were available, it was not straightforward to find the optimal structure. Our finding of the induced-fit conformation via REMD widened the spectrum of REMD applications in the drug design field by obtaining various sampling spaces. Although the structure identified here showed a small scale change with mainly adjustment of the Phe93 in the binding pocket, many residues at the interface, including the N-terminal and switch-loop, were also altered. These collective rearrangements resulted in the ligand-induced-fit like conformation at the interface.

We used an ensemble-docking strategy to account for receptor flexibility. Totrov and Abagyan proposed an efficient method of generating multiple receptor conformations with limited flexibility for the side-chains of various conformations of 6–12 residues.<sup>20</sup> They generated a set of multiple receptor conformations from the NTP-MD ensemble structures, normal mode representative structures, and the NMR-ensemble structure. In addition to those ensemble structures, the REMD sampling structures can be additionally served as enhanced ensemble receptor conformations. Another advantage of REMD sampling method is that it allows the utilization of the entire sampled space by constructing a PMF to extract free energy information (e.g., Figure 4).

**Binding Free Energy Gain to Induce Receptor Conformation along the Reaction Coordinate.** In the series of energy results shown in Figure 2A (potential energy of refined MD), Figure 6 (morphed binding energy), and Figure 7B (PMF profile from US), the potential energy of the best-REMD snapshot was observed to be stabilized upon binding of the Abbott8 compound, which indicated that the ligand induced conformational stabilization. The energy gained from the binding of the Abbott8 compound was used to rearrange the conformation toward the induced conformation. The most probable potential energy of the refined MD was lower than the induced-fit structure initially found with REMD in Figure 2A, which is consistent with the stabilization of binding energy shown in Figure 6 and the energy barrier seen in Figure 7B. Around a 10 Å distance of reaction coordinate, the seventh morphed binding energy began to stabilize (Figure 6); at the same distance, the corresponding receptor conformation is in the transition state of the PMF profile (Figure 7B). Once the ligand bound to the optimally arranged survivin dimer conformation at 10 Å, the gained binding energy was used to adjust residues of receptor to further stabilize it in a conformation suitable for ligand binding.<sup>21</sup> Consequently, the PMF showed a local energy minimum around 11.5 Å (Figure 7B).

The mitosis/apoptosis-switch loop is predicted to be a functionally important region via CPC X-ray structure, where Borealin mimics the survivin counter monomer structure at the interface.<sup>3</sup> This switch loop region was also identified as critical in this simulation study, where it served as a crucial reaction coordinate that captured the conformational change from the X-ray structure toward the structure suitable for receptor binding.

**Comments on Free Energy Calculation.** Free energy calculation remains a very difficult challenge for computational chemistry. Here the calculated binding free energy being close to experimental value may be a special case. In MMPB(GB)SA end-point evaluation method, a large set of model assumptions has been made and the results are very system dependent.<sup>23</sup> However, in this case we have some confidence. (1) The system

is relatively small, entropy estimation via normal mode approximation seems valid. (2) For desolvation free energy evaluation with small and globular protein, a neutral ligand with limited polarization and most up-to-date OBC GB (Generalized Born) model (Amber igb=5<sup>24</sup>), the estimated Born radii and the electrostatic contribution of ligand desolvation could be reliable. For barrier calculation, we used two well tested conventional free energy estimators, MBAR and WHAM, to cross-validate each other, resulting in identical PMF profiles with low standard deviations.

**The “Population-Shift” versus “Induced-Fit” Debate.** In protein–ligand recognition, the theory of population shift claims that the protein conformations pre-exist, which makes sense according to the free energy landscape theory. The binding process is conformation selection and thermodynamic re-equilibration. The Abbott8-survivin binding simulation actually supports both population-shift and induced-fit claims. The complexed survivin conformation does pre-exist as this conformation can be detected via REMD generalized ensemble sampling of survivin structure alone. The free energy of this conformation is about 3–4 kcal/mol higher than that of the conformation determined via X-ray crystallography, which translates into ~0.4% population at 300 K. However, the binding process is a strikingly “induced-fit” dynamic. The hydrogen-bond coupled electrostatic steering between survivin and Abbott8 seems crucial for the complexation to overcome the steric blockage by the switching loop, at least initially. Mere thermomotions are not enough as demonstrated by the sampling failure of regular NTP-MD simulation.

**General View for the Simulation Protocol Setup.** REMD was originally intended for sampling a complete thermodynamic ensemble, the so-called “converged” sampling, so it may be used to study protein folding/unfolding, for example. Indeed, initially we had experimented with a much larger temperature range of 280–600 K for a total of 98 ns to yield a bigger ensemble including unfolded conformations, and we docked Abbott8 into that ensemble and found the correct survivin/Abbott8 complex consistent with NMR data also. However, a “fully converged” ensemble is not necessary in our opinion, at least in this case. We observed that at 340 K, the survivin dimer interface started to dissociate significantly. In this case, we know that interface-perturbing Abbott8 is not able to cause dimer dissociation, and that is why we chose 340 K as our upper limit for REMD. For temperature interval, we followed conventional setup to choose the number of replicas described in the method section and then tried both evenly distributed and exponentially distributed intervals; in this case, the evenly distributed setup yields better exchange ratio among replicas. Since the temperature range and number of conformations are relatively moderate, a prefiltering and clustering before docking screening was not needed. In conclusion, depending on the molecular systems and simulation goals, different protocol setups need to be explored in applying this new protein/ligand recognition simulation methodology.

## Concluding Remarks

We have identified the optimal conformation suitable for Abbott8 compound binding to the survivin dimerization interface via REMD and ensemble dockings. In the optimal conformation, the F93–F101 distance (reaction coordinate) was opened to 12.5 from 4.5 Å. The ensuing free energy profile analysis for the dynamic binding process reveals a free energy barrier of 3–4 kcal/mol and the induced-fit happens due to the binding energy of Abbott8, primarily through hydrogen-bond coupled electrostatic steering. Ultimate dimer-

disrupting compounds may be discovered via computational design based on the optimal receptor structure found in this study. In general, the combination of REMD generalized ensemble sampling with ensemble docking and free energy pathway analysis may provide a novel research protocol for the simulation of protein–ligand induced-fit recognition.

**Acknowledgment.** We gratefully thank the Ohio Supercomputer Center (OSC) for the donation of computing time on the Glenn cluster. This work was supported by the Faculty startup fund from the College of Pharmacy to C.L.

**Supporting Information Available:** The detailed normal-mode analysis and PMF path construction are described. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References and Notes

- (1) Chantalat, L.; Skoufias, D. A.; Kleman, J. P.; Jung, B.; Dideberg, O.; Margolis, R. L. Crystal structure of human survivin reveals a bow tie-shaped dimer with two unusual alpha-helical extensions. *Mol. Cell* **2000**, *6*, 183–189.
- (2) Shi, Y. Survivin structure: crystal unclear. *Nat. Struct. Biol.* **2000**, *7*, 620–623.
- (3) Ruchaud, S.; Carmena, M.; Earnshaw, W. C. The chromosomal passenger complex: one for all and all for one. *Cell* **2007**, *131*, 230–231.
- (4) Jayaprakash, A. A.; Klein, U. R.; Lindner, D.; Ebert, J.; Nigg, E. A.; Conti, E. Structure of a Survivin-Borealin-INCENP core complex reveals how chromosomal passengers travel together. *Cell* **2007**, *131*, 271–285.
- (5) Wendt, M. D.; Sun, C.; Kunzer, A.; Sauer, D.; Sarris, K.; Hoff, E.; Yu, L.; Nettesheim, D. G.; Chen, J.; Jin, S.; Comess, K. M.; Fan, Y.; Anderson, S. N.; Isaac, B.; Olejniczak, E. T.; Hajduk, P. J.; Rosenberg, S. H.; Elmore, S. W. Discovery of a novel small molecule binding site of human survivin. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 3122–3129.
- (6) Altieri, D. C. Survivin, cancer networks and pathway-directed drug discovery. *Nat. Rev. Cancer* **2007**, *8*, 61–70.
- (7) Pennati, M.; Folini, M.; Zaffaroni, N. Targeting survivin in cancer therapy. *Expert Opin. Ther. Targets* **2008**, *12*, 463–476.
- (8) Call, J. A.; Eckhardt, S. G.; Camidge, D. R. Targeted manipulation of apoptosis in cancer treatment. *Lancet Oncol.* **2008**, *9*, 1002–1011.
- (9) Sun, C.; Nettesheim, D.; Liu, Z.; Olejniczak, E. T. Solution structure of human survivin and its binding interface with Smac/Diablo. *Biochemistry* **2005**, *44*, 11–17.
- (10) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolosky, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seitun, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*, University of California: San Francisco, 2008.
- (11) Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.* **2002**, *116*, 9058–9067.
- (12) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
- (13) *Glide*, version 5.0; Schrödinger, LLC: New York, 2008.
- (14) R Development Core Team. *R: A language and environment for statistical computing*; R Foundation for Statistical Computing: Vienna, Austria, 2005; ISBN 3-900051-07-0, <http://www.R-project.org> (accessed 2006).
- (15) Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
- (16) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.
- (17) Grossfield, A. WHAM, version 2.0.2, 2008. <http://membrane.urmc.rochester.edu/Software/WHAM/WHAM.html> (accessed 2008).
- (18) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389.
- (19) Okur, A.; Wickstrom, L.; Layten, M.; Geney, R.; Song, K.; Hornak, V.; Simmerling, C. Improved Efficiency of Replica Exchange Simulations through Use of a Hybrid Explicit/Implicit Solvation Model. *J. Comput. Chem.* **2006**, *2*, 420–433.
- (20) Totrov, M.; Abagyan, R. Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178–184.
- (21) Fersht, A. *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*; W.H. Freeman: New York, 1999; Chapter 10.

- (22) Mongan, J. Interactive essential dynamics. *Comput.-Aided Mol. Des.* **2004**, *18*, 433–436.
- (23) Mobley, D. L.; Dill, K. A. Binding of Small-Molecule Ligands to Proteins: “What You See” Is Not Always “What You Get”. *Structure* **2009**, *17*, 489–98.
- (24) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins* **2004**, *55*, 383–394.

JP911085D