JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Link between Allosteric Signal Transduction and Functional Dynamics in a Multisubunit Enzyme: *S*-Adenosylhomocysteine Hydrolase

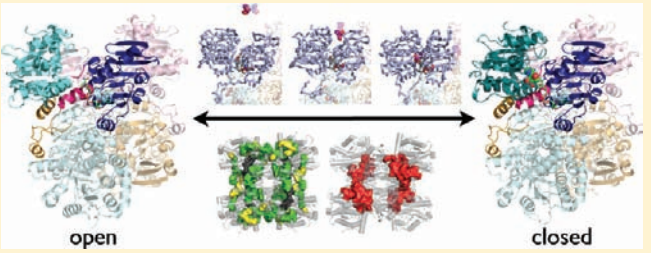
Yoonji Lee,[†] Lak Shin Jeong,[†,‡] Sun Choi,*[,†] and Changbong Hyeon*[,¶]

[†]College of Pharmacy, Division of Life and Pharmaceutical Sciences and National Core Research Center for Cell Signaling and Drug Discovery Research, [‡]Department of Bioinspired Sciences, Ewha Womans University, Seoul 120-750

[¶]School of Computational Sciences, Korea Institute for Advanced Study, Seoul 130-722, Republic of Korea

Supporting Information

ABSTRACT: *S*-adenosylhomocysteine hydrolase (SAHH), a cellular enzyme that plays a key role in methylation reactions including those required for maturation of viral mRNA, is an important drug target in the discovery of antiviral agents. While targeting the active site is a straightforward strategy of enzyme inhibition, evidence of allosteric modulation of active site in many enzymes underscores the molecular origin of signal transduction. Information of co-evolving sequences in SAHH family and the key residues for functional dynamics that can be identified using native topology of the enzyme provide glimpses into how the allosteric signaling network, dispersed over the molecular structure, coordinates intra- and intersubunit conformational dynamics. To study the link between the allosteric communication and functional dynamics of SAHHs, we performed Brownian dynamics simulations by building a coarse-grained model based on the holo and ligand-bound structures. The simulations of ligand-induced transition revealed that the signal of intrasubunit closure dynamics is transmitted to form intersubunit contacts, which in turn invoke a precise alignment of active site, followed by the dimer−dimer rotation that compacts the whole tetrameric structure. Further analyses of SAHH dynamics associated with ligand binding provided evidence of both induced fit and population shift mechanisms and also showed that the transition-state ensemble is akin to the ligand-bound state. Besides the formation of enzyme-ligand contacts at the active site, the allosteric couplings from the residues distal to the active site are vital to the enzymatic function.



INTRODUCTION

S-adenosylhomocysteine (SAH) hydrolase catalyzes the hydrolytic cleavage of SAH to adenosine and L-homocysteine. Inhibition of this enzyme causes the accumulation of SAH and consequently suppresses *S*-adenosyl-L-methionine dependent transmethylation via a feedback inhibition mechanism. Since the methylation at the 5′-terminus of mRNA is crucial for the viral replication, SAHH is a promising target for the discovery of broad spectrum antiviral agents.[1]

Developing therapeutic agents that can directly bind and regulate the active site of a biological target has been a dominant pharmacological strategy.[2,3] For SAHH, various adenosine analogues, including carbocyclic adenosine, neplanocin A, 3-deazaneplanocin A,[1] and fluoro-neplanocin A (F-NpcA),[4] are the known inhibitors that directly target the active site.[4] However, conformational flexibility of enzyme structures gleaned in X-ray and NMR experiments and the presence of allosteric site revealed in mutational studies highlight the allosteric couplings of residues distal to the active site as an another important principle in drug design strategy.[5−7] Although the formation of specific enzyme−substrate contacts in the active site is required for the catalytic activity, allosteric orchestration among residues, dispersed over the molecular architecture, is also essential to regulate conformational fluctuations, so as to assist a precise positioning of catalytic elements in the active site.[8] SAHH, an enzyme consisting of chemically identical four subunits, each of which undergoes the open-to-closed (O→C) transition in response to substrate binding, is an interesting system to study the link between enzymatic function and allosteric dynamics beyond monomeric enzyme.[9−15]

Structure, dynamics, and catalytic function are the three main themes in understanding enzymes.[6,9,16−20] For the past decades, much effort has been made to decode the link between the allosteric signaling of enzymes, conformational dynamics, and their function by using both theories[10,14,21−24] and experiments[13,20,25−27] and has recently been extended to the studies of molecular motors.[28−30] To gain microscopic understanding to the allostery in SAHH and its implication to the catalysis, we employed multifaceted computational approaches: (i) Statistical coupling analysis (SCA)[21,31,32] was used to reveal networks of co-evolving amino acid residues in the SAHH family; (ii) structural perturbation method (SPM) adapting normal mode analysis (NMA) identified the network of hot spot residues associated

Received: July 15, 2011
Published: October 24, 2011

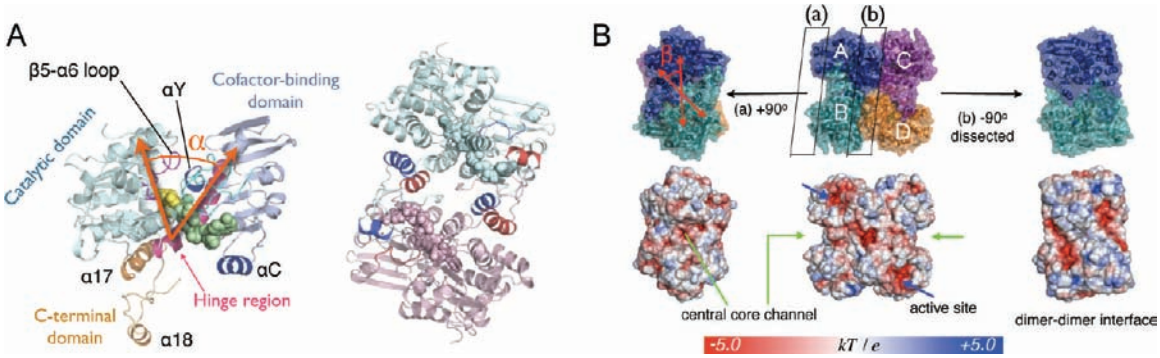


Figure 1. Structure of SAHH. (A) A subunit structure consisting of the catalytic (residues 1−181 and 355−385), cofactor-binding (residues 197−351), and C-terminal domains (residues 386−432) are shown on the left. The angle α is defined among W112, K186, and W310. A dimer structure assembled by exchanging the C-terminal domains is shown on the right with structural motifs associated with the dimer interactions being colored in blue and red. (B) Homotetrameric structure. Each monomer is represented by blue, cyan, purple, and orange surfaces. β is the angle between the two vectors, one from CM_A to CM_B and the other from CM_C to CM_D, where CM_X is the center of position of a subunit X. Electrostatic potentials on the molecular surface calculated by solving Poisson−Boltzmann equation at 0.2 M salt condition show negatively charged central core channel and active site. The side view of the structure and the dimer−dimer interface are shown on the left and right, respectively.

with the functional motion of SAHH;[23] and (iii) while the SCA and SPM provide hints as to the correlation between allosteric signaling network and conformational dynamics, it is difficult to gain from these two static analyses further insights into a large-scale conformational change, such as a shifting of statistical ensemble and heterogeneity in the transition routes. To this end, we performed Brownian dynamics (BD) simulations of the O→C transition of tetrameric SAHH in response to ligand binding by using a structure-based coarse-grained model.[10,12,30,35−37]

Among the various stages of the SAHH enzymatic cycle that involve multiple catalytic processes,[33,34] which occur on time scales of ∼10−100 ms, the main focus of our simulation is on the fast ∼1−100 μs conformational dynamics of SAHH before and after as well as in the process of substrate binding. Despite the large time scale gap between catalysis and conformational dynamics associated with ligand binding, it is important to understand the conformational fluctuations and dynamics, since they lie at the core of allosteric regulation of catalytic power in enzymes. In this paper, by analyzing the dynamics resulting from our simulation, we highlight the link between the functional dynamics and the allosteric signaling implicated in the SCA and SPM.

METHODS

SAHH Structure. A subunit structure, consisting of 432 residues, retains catalytic, cofactor-binding and C-terminal domains. An α helix (αA1) and loops (residues 182−196, 352−354) join the catalytic domain with the cofactor-binding domain, forming a hinge region (Figure 1A, left, see also Figure S1, Supporting Information). The interaction between C-terminal domain of a subunit and cofactor-binding domain of the partner subunit (Figure 1A, right) is employed to assemble two dimer complexes (AB and CD) into the tetramer. A negatively charged central core channel constitutes the interface between A and B (or C and D) (Figure 1B). A further assembly between the two dimers (AB and CD), whose interface exhibits electrostatic complimentarity, forges an oblate-like tetrameric structure (Figure 1B). Depending on the presence of ligand in the binding cleft between catalytic and cofactor-binding domains, SAHHs select either open or closed conformation whose root-mean-square deviation (rmsd, Δ_{OC}) is 4.2 Å. The difference of the two conformations can also be quantified using an angle, α, defined among W112, K186, and W310, which is $\alpha = 70°$ for the open and $\alpha = 50°$ for the closed structures Figure 1A. In addition, an angle β measures the relative orientation between AB and CD dimers; $\beta = 3°$ for the open and $\beta = 15°$ for the closed structures (Figure 1B).

SCA on SAHH Family. If two amino acid residues in a protein are functionally or energetically coupled, a mutation at one site is expected to influence other sites to level off the energetic perturbation, so that the protein can retain its functional state. While double or triple mutant experiment is laborious to prove this proposition, evolutionary data imprinted in multiple sequence alignment (MSA) of a protein family still allow one to infer the coupling between two distal sites. SCA is a method to identify a set of co-evolving amino acid sites by performing a mutation "experiment" on the MSA of a protein family.[21,31,32]

The whole sequence space of SAHH family, which includes 1156 sequences over the entire organisms, was clustered by their sequence identities, and 252 sequences under 75% identities were finally selected. This diverse sequences of SAHH family were used to produce MSA using the multiple sequence comparison by the log-expectation (MUSCLE) program.[38] Each of the sequences has 1242 positions, including gaps in the full MSA. Using MSA of the SAHH family, we calculated the statistical free energy-like function defined as $\Delta G_i/k_BT^* = ((1/C_i)\Sigma_{\alpha=1}^{20}[p_i^\alpha \log(p_i^\alpha/p_\alpha)]^2)^{1/2}$, where C_i is the number of types of amino acid at position i along the sequence, α denotes amino acid species, p_i^α is the frequency of an amino acid α at position i, and p_α is the frequency of an amino acid α in the full MSA. ΔG_i (Figure 2A) identifies highly conserved sequence positions in ligand- and cofactor-binding site and the C-terminal domain that constitutes a part of the cofactor-binding site (Figure 2B).

Further analysis of the MSA enabled us to identify the set of mutually correlated residues in SAHH from an evolutionary perspective. The SCA[21,31,32] calculates $\Delta\Delta G_{ij}/k_BT^* = ((1/C_i)\Sigma_{\alpha=1}^{20}[p_{i,R}^\alpha \log(p_{i,R}^\alpha/p_\alpha) - p_i^\alpha \log(p_i^\alpha/p_\alpha)]^2)^{1/2}$, where $p_{i,R}^\alpha$ is the frequency of amino acid in a restricted subalignment with respect to the position j, a score matrix that quantifies the effect of sequence "perturbation" at position j on position i by restricting the full MSA into a subalignment with the most conserved sequence at the j-th position and quantifying the sequence variation at the i-th position of subalignment with respect to the same position of the full MSA. Since the choice of appropriate size for subalignments (f) is critical in obtaining statistically significant correlations between residues in a protein family, the smallest value of f should be chosen to satisfy the central limit theorem.[32] In order to determine the suitable f value, we built subalignment from the MSA with different f (<1) values by randomly choosing fN_{MSA} sequences, where N_{MSA} is the size of the full

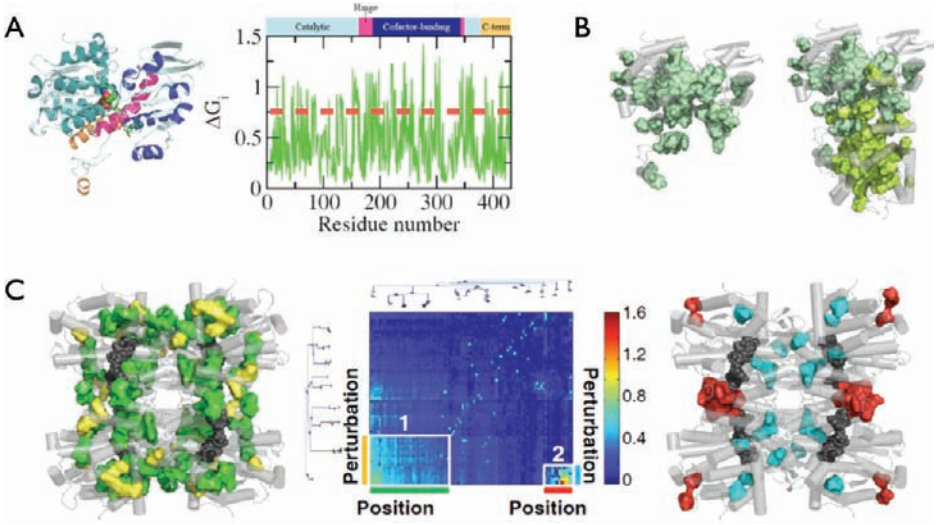


Figure 2. Statistical coupling analysis of SAHHs. (A) $\Delta G_i/k_BT^*$ value of the i-th residue in SAHH monomer. The catalytic (cyan), hinge (magenta), cofactor-binding (dark blue), and C-terminal (orange) domains are depicted in different colors. (B) Highly conserved residues with $\Delta G_i > 0.75$ from A are mapped on the monomer (left) and dimer (right) structures. (C) Reordered $\Delta\Delta G_{ij}$ matrix that identifies the two clusters (white boxes) is shown in the middle, and the corresponding two co-evolving clusters of sequences are depicted on the SAHH structure (see Supporting Information). For cluster1, perturbation and position clusters are colored in yellow and green, respectively. For cluster2, perturbation and position clusters are colored in cyan and red, respectively.

MSA. For a given f, we generated 1000 subalignments and computed $\langle \Delta G_i^s \rangle_f = \frac{1}{1000}\sum_{k=1}^{1000} \overline{\Delta G_i^s}$ and $\sigma_f^2 = (\Delta G_i^s)^2 - \overline{\Delta G_i^s}^2$, where $\overline{\Delta G_i^s}$ means the average ΔG_i values in a randomly chosen subalignment. By analyzing the distribution of these values, we selected $f = 0.5$ to statistically significant subalignment size (Figure S2A, Supporting Information), and 213 positions were allowed to be perturbed in this condition. In the $\Delta\Delta G_{ij}/k_BT^*$ scoring matrix, $\Delta\Delta G_{ij}$ with a high value signifies that the i-th position is susceptible to the perturbation at the j-th position. After filtering out the noise, the allosteric signaling network was obtained by adopting the coupled two-way clustering algorithm to cluster the position and perturbation elements of $\Delta\Delta G_{ij}$ exhibiting a high signal.[39,40]

Elastic Network Model (ENM) and Structural Perturbation Method (SPM). In ENM, a protein structure is represented as a network of beads connected with mechanical springs of uniform force constant k_0. The potential energy for ENM is given as $E_{ENM} = (1/2)\Sigma_{ij} k_0(r_{ij} - r_{ij}^o)^2 \Theta(r_{ij}^o - R_c)$, where r_{ij} is the distance between residues i and j, and $\Theta(...)$ is the heaviside step function, which connects two coarse-grained centers (C_α atoms) if the distance r_{ij}^o in the native structure is within a cutoff distance R_c. By calculating the Hessian matrix of E_{ENM}, one can obtain normal modes ($\vec{v}_M$, $M \geq 7$, $1 \leq M \leq 6$ are associated with translational and rotational degrees of freedom) of a protein architecture. Given the open ($\vec{r}_O$) and closed structures ($\vec{r}_C$) of SAHH, the overlap between $\vec{v}_M$ and the O→C conformational change ($\vec{r}_{O\to C}$) is evaluated by calculating cos ($\vec{v}_M \cdot \vec{r}_{O\to C}$) (Figure 3).

The SPM assesses the importance of a residue in the elastic network by locally perturbing the residue and probing its response. The perturbation is realized by changing the force constant of the springs that link the residue and its neighbors. The response to the perturbation in the frequency of mode M, which overlaps the most with the O→C structural transition, is calculated using $\delta\omega(M,n) = \vec{v}_M^T \cdot \delta H \cdot \vec{v}_M$, where δH is the Hessian matrix of the following perturbed energy potential: $\delta E_{ENM} = (1/2)\Sigma_{ij}\delta k_0(r_{ij} - r_{ij}^o)^2 \Theta(r_{ij}^o - R_c)$. The higher $\delta\omega(M,i)$ signifies an importance of i-th residue with respect to the mode M.

Energy Potential and Simulations. To define the energy potential of SAHH, we used an X-ray crystal structure of human SAHH (PDB id: 3NJ4) that contains F-NpcA as the closed structure and a homology model based on rat SAHH as the open (or holo) structure.[4] The open structure of SAHH should be stable in the absence of ligand, but the formation of the inter-residue contacts incorporated from the closed structure should bring SAHH from the open to closed forms in the presence of SAHH−ligand interactions, by further stabilizing the ligand-bound SAHH structure.[12] To implement this dynamical behavior, we used the self-organized polymer (SOP) energy potential:[10,12,36] $H_{tot} = H^{SAHH} + H^{lig} + H^{SAHH-lig}$ where $H^{SAHH} = \Sigma_{i=A,B,C,D} H_{intra}^{SAHH}(i) + \Sigma_{i>j} H_{inter}^{SAHH}(i,j)$, and $H^{SAHH} = H_{FENE} + H_{nb}$ consists of the terms involving intrasubunit connectivity using finite extensible nonlinear elastic potential (H_{FENE}) and nonbonded potential term (H_{nb}) that stabilizes either one of the open or closed monomer structures depending on the presence of ligand at its binding cleft. The energy contributions from the H_{nb} are the sum of open native contacts ($H_{nb}^{(O)}$) and pure closed native contacts $H_{nb}^{(C\cap O^c)}$.[12] The enzyme−ligand interaction ($H^{SAHH-lig}$) should contribute to the SAHH compaction.

We simulated the SAHH conformational dynamics under an overdamped condition by integrating the Brownian dynamics algorithm.[12,41] The ligand binding dynamics were simulated by releasing four ligands at ≈20 Å away from each binding cleft under a periodic and weak harmonic constraint ($k = 0.035$ pN/nm) imposed every 25 μs, which was used to ensure the ligands to bind the binding cleft within our simulation time. The initial distance of the ligands from each binding pocket provides enough time to randomize the orientation of the ligand before it reaches the binding cleft (see also Movies S1 and S2, Supporting Information). While the form of SOP potential and simulation strategy described above is essentially the same as those used in the previous study on protein kinase A,[12] the additional term that takes into account the intersubunit interaction H_{inter}^{SAHH} substantially enriches the resulting dynamics.

RESULTS AND DISCUSSION

Allosteric Signaling Network Inferred from Co-evolving Residues. Employing SCA to the MSA of SAHH family, we identified two co-evolving clusters of residues (Figure 2C and Table 1 for the list). It is noteworthy that since SAHHs function

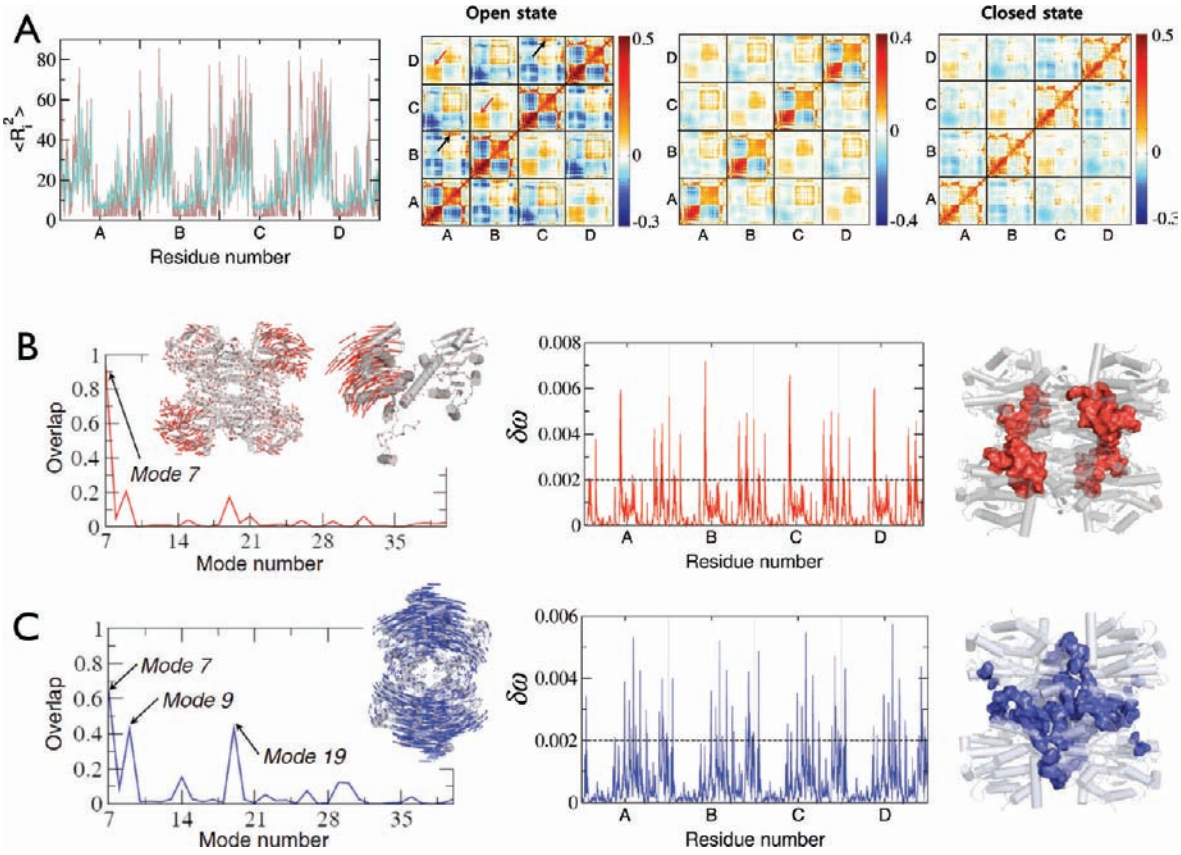


Figure 3. Normal mode analysis and structural perturbation method applied to SAHH. (A, left) Comparison between mean square fluctuation (cyan) and crystallographic b-factor (brown) of the open form of rat SAHH (PDB code: 1B3R). (A, right) Covariance matrices of SAHH open (left) and closed (right) forms. The differences between the two matrices are shown in the middle. Our calculations using ENM compare well with those in a previous study.[46] (B) Overlap between $\vec{r}_{OC}$ and normal mode M ($\vec{v}_M$) of the open structure. Maximally overlapping normal mode ($M = 7$) are displayed using lines in the structures (left). The effect of structural perturbation, $\delta\omega(M,i)$, is calculated for $\vec{v}_{M=7}$ (middle). The residues with a high $\delta\omega$ value with respect to $\vec{v}_{M=7}$ are depicted in the structure on the right and listed in the Table S1, Supporting Information. (C) The same calculation for the closed structure by using $\vec{v}_{M=7}$ as the maximally overlapping mode with $\vec{r}_{OC}$.

Table 1. Evolutionarily Covarying Residues Identified by Statistical Coupling Analysis (SCA)

	cluster 1	cluster 2
perturbations	19 62 65 79[a,b] 81[a] 92[a]	81 92
	103 136 195[a] 196[a] 224[a]	195 196
	293 359[a] 386[a] 390 392	293 359
positions	24 56 59 76 78 79[a,b]	122 123
	81[a] 82 87 88 92[a] 179	125 425
	183 192 195[a] 196[a] 197 208	426[b] 427
	223 224[a] 248 251 322 328	428 429
	349 359[a] 368 369 386[a] 398	430[b] 431
	406 407 420 424	432[b]

[a] Residues identified in both the perturbation and position clusters. [b] Highly conserved residues.

as a tetramer, the co-evolving clusters revealed from monomer sequence are indeed the consequence of allosteric communication between residues spread over the entire tetrameric architecture. Most of co-evolving residues identified by the SCA are not well-conserved over the SAHH family except for the residues in C-terminal domain and the C79 that is in direct contact with the ligand (Figure S2C, Supporting Information): (i) The cluster 1 shows that a number of residues in green clusters around the active site are susceptible to the perturbation of residues in yellow clusters. Given that the active site is the core region for SAHH function, it is not surprising that the residues susceptible to perturbations are distributed around the binding cleft. Moreover, the region of central core channel is also detected to be a part of the main allosteric signaling network, which includes C195, R196, and E197. In particular, C195 is known to be a key residue for the catalysis and the maintenance of the reduction potential at the 3′-position.[42] (ii) The cluster 2 includes two spatially disconnected sets of residues in the C-terminal domain (from F425 to Y432) and the residues (G122, P123, and D125) in the catalytic domain that are located away from the active site. A mutation in "cyan" clusters perturbs the residues in "red" clusters. According to the mutation studies, K426 and Y430 in this region are the critical residues that affect the cofactor affinity and/or the assembly of tetrameric structure.[43,44] Furthermore, although it was not detected in the SCA, a mutation of R49 in direct contact with D125 that belongs to the cluster 2 is reported to dramatically reduce SAHH activity.[45]

Network of Hot Spot Residues Responsible for the Collective Dynamics of SAHH. Normal mode analysis (NMA) based on the ENM was used to grasp the characteristics of SAHH dynamics evolving from the native topology. Along with the

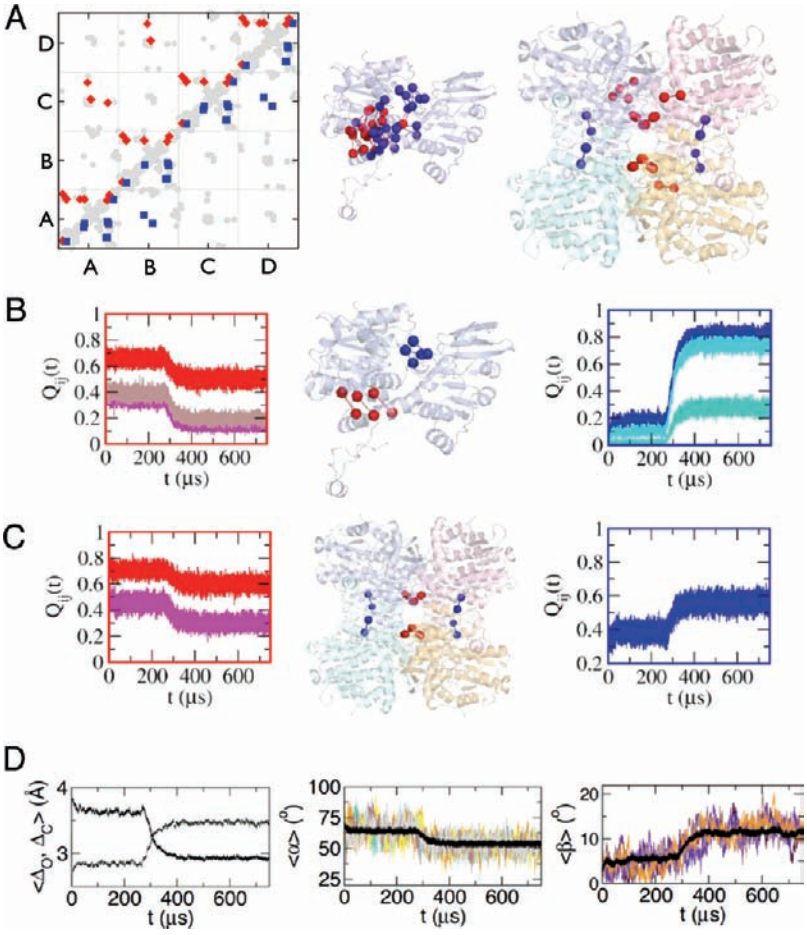


Figure 4. BD simulations of the O→C transitions in SAHH. (A) Native contact probability (Q_{ij}) changes as a result of the conformational change in response to the ligand binding. Among the native contact pairs, Q_{ij} are colored in blue (red) when Q_{ij} increases (decreases) more than 10%. The position of corresponding residue pairs in a SAHH subunit and at the subunit interface is depicted in the structures on the right. (B and C) Representative time traces of native contacts $Q_{ij}(t)$ for the intra- (B) and intersubunit (C) residue pairs. The residue pairs with $\delta Q_{ij}(t) < 0$ after ligand binding are marked in red, and those with $\delta Q_{ij}(t) > 0$ are in blue; 189−361 (red), 185−364 (magenta), 367−394 (brown), 83−347 (blue), 83−348 (cyan), and 82−343 (turquoise); AC(BD):231−235 (red), AC(BD):234−235 (magenta), and AB(CD):185−247 (blue). (D) Time-dependent changes of rmsd relative to the open and closed forms ($\langle\Delta_O\rangle$ and $\langle\Delta_C\rangle$) and $\langle\alpha\rangle$ and $\langle\beta\rangle$ are shown from left to right.

previous studies,[46,47] the covariance matrices calculated for the open and closed forms of SAHH show how the enzyme fluctuates in each form (see Figure 3A). The catalytic domain is more mobile than the cofactor-binding domain, and the overall flexibility decreases in the closed conformation. They also indicate that the cofactor-binding domain and the C-terminal domain of the partner subunit are cross-correlated (see black arrows), in consistent with the structural organization of SAHH dimer where the C-terminal domains are exchanged between the adjacent subunits. Interestingly, the catalytic domain of the A (or B) subunit is cross-correlated with that of the D (or C) subunit (see red arrows). The largest overlap of normal mode with the O→C conformational change ($\vec{r}_{O\rightarrow C}$) (see Methods) is found with the lowest frequency normal mode (mode 7) in both the open and closed forms [Figure 3B (left) and C (left)]. The closure motion of catalytic domain relative to the cofactor-binding domain and the relative rotation between AB and CD dimers are the main fluctuations characterizing the global dynamics of open and closed forms of SAHH.

To determine "hot spot" residues controlling the O→C conformational dynamics, we employed the SPM,[23] which complements the results of SCA. The SPM mimics the point mutation of an enzyme by increasing the spring constant of a residue in the elastic network. Perturbation to a residue important for functional motion leads to a large change in the normal-mode frequency (high $\delta\omega$, see Methods Section). For the closure dynamics described by mode 7 of the open SAHH structure, key residues with high $\delta\omega$ are mainly distributed in the hinge region (Figure 3B, right). Notably, this finding conforms to the recent time-resolved fluorescence anisotropy measurements that probed the domain motions of mutant SAHH, where importance of residues M351, H353, and P354 at the hinge region was highlighted.[48] In contrast, for the dimer−dimer rotation dynamics described by the mode 7 of the closed structure, high $\delta\omega$ residues are found at the interfaces between subunits (Figure 3C, right). Key residues controlling the next higher frequency modes are distributed mainly at the interfaces between the subunits, contributing to the concerted intersubunit dynamics (see Figure S3, Supporting Information, for mode 9).

Global Motion of SAHH during the O→C Transitions upon Ligand Binding. The conformational dynamics and its dynamical response to ligand binding are prerequisite to understanding the enzyme function. To this end, we performed BD simulations using a coarse-grained model of the whole tetrameric structures

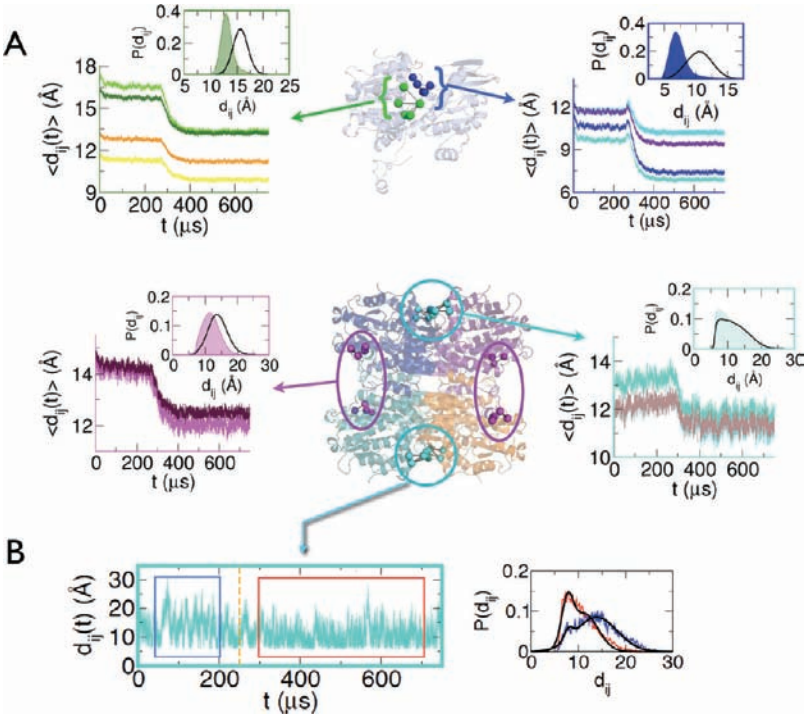


Figure 5. Dynamics of intra- and intersubunit residue pairs. (A) Intra- and intersubunit dynamics averaged over the ensemble of trajectories probed by $\langle d_{ij} \rangle$ and the corresponding histogram, $P(d_{ij})$, before (black line) and after (colored area) the ligand binding. The intrasubunit residue pairs consisting of the catalytic site exhibit monotonic (green spheres) and nonmonotonic decays (blue spheres). The intersubunit residue pairs at AB (or CD) and AC (or BD) interfaces are displayed in purple and cyan spheres, respectively. The list of residue pair dynamics depicted in each panel is: 79−301, 131−301*, 157−301, and 159−301 for green spheres; 83−347, 83−348* 83−351, and 85−351 for blue spheres; AB(CD) 166−418* and AB(CD) 162−416 for magenta spheres; AC(BD) 17−321*, 17−322, and 20−321 for cyan spheres, where anterisk (*) denotes the residue pairs whose histogram is shown in the insets. (B) An individual time trace of $d_{\mathrm{AC(BD):W17-I321}}(t)$ at the AC subunit interface displays a biomodal hopping transition, and its population change is shown in the right panel. The histograms before and after the ligand binding (blue and red lines, respectively) fitted to a double Gaussian function (black lines) show that there is a "shifting of population" from $d_{\mathrm{AC(BD):W17-I321}} \approx 14$ to 8 Å.

in the holo and ligand-bound states (see Methods Section). Each trajectory consists of 750 μs simulations; the first 250 μs simulates the equilibrium stage of SAHH in the absence of ligand, and the next 500 μs is for the O→C transition dynamics in response to the ligand binding. We probed the ligand binding by using the distance between the fluorine atom of F-NpcA and C4 atom of the nicotinamide ring of the cofactor NADH (Figure S4, Supporting Information). Among the total 60 simulated trajectories, we observed complete bindings of four ligands in 39 trajectories during our simulation time. A comparison of the fraction of native contact map (Q_{ij}) before and after the ligand binding revealed that Q_{ij} increases and decreases depending on residue pairs (Figure 4A), suggesting that the structural reorganization due to the O→C transition involves formation and breakage of multiple residue pairs (Figure 4A, B, and C). The Q_{ij} between the β5−α6 loop and αY helix (marked with blue spheres in Figure 4B), located at the entrance of the binding cleft, dramatically increases after the ligand binding, while the Q_{ij} between αA1, α17, and α15 (marked with red spheres in Figure 4B) decreases. In case of the intersubunit contacts, key increases of residue contacts are made between the hinge region and the αC helix of the partner subunit (Figure 4C).

Prior to the ligand binding, the enzyme structure remains largely in the open state as expected. Nevertheless, the fluctuation of the rmsd relative to the closed structure (Δ_C) occasionally brings SAHH structure to $\Delta_C < 2.5$ Å even without ligand in the binding cleft (Figure 7B). As a result of the ligand binding, the Δ_C decays from 3.7 to 2.9 Å in 34 μs (Figure 4D left), the angle α decreases from 69 to 53° (Figure 4D middle), and the angle β changes from 5 to ≈12° (Figure 4D, right).

Kinetic Hierarchy of the O→C Transition. The inter-residue distance between i-th and j-th residues, $\langle d_{ij}(t) \rangle$, where $\langle ... \rangle$ denotes an ensemble average over trajectories, probes microscopic details of O→C transitions described above. Histograms of d_{ij}, $P(d_{ij})$, before and after the ligand binding quantify the change in average and fluctuation of inter-residue distance (Figure 4). The average transition time for each residue pair, obtained from multi (mostly single) exponential fit to the $\langle d_{ij}(t) \rangle$, reveals the order of events that occur during the O→C transition dynamics upon ligand binding (Figure 6).

Closure Dynamics of the Binding Cleft. The structural motifs in the binding cleft, composed of several loops and an α helix, directly interact with a ligand that mediates the association between the catalytic and cofactor-binding domains (Figure 1A). We employed several residue pairs in these loops as reporters of intrasubunit closure dynamics. The dynamics of residue pairs associated with G277, H301, A340, and R343 from three representative loops consisting of binding cleft show single exponential decays with time constants of $\tau = 35-38$ μs (Figure 5A, green panel). The ligand binding reduces not only the average distances between the catalytic and cofactor-binding domains but also their fluctuations (see the inset of green panel in Figure 5A for D131−H301 pair). These results are in good agreement with the experimental data of fluorescence anisotropy measurements,[49]

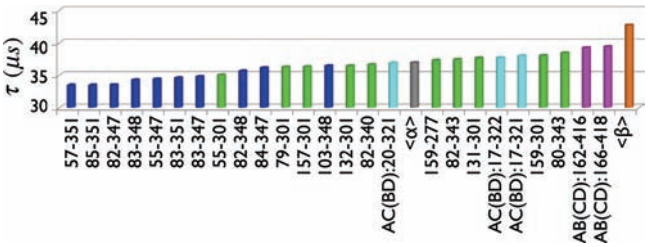


Figure 6. Kinetic hierarchy revealed in the SAHH dynamics. Average transition times for various residue pairs discussed in Figure 5 and α and β angle dynamics are shown in an increasing order. The colors of bar graphs are identical to those used in the previous figures.

confirming that the intrasubunit fluctuation is substantially reduced after the ligand binding. Among the residue pairs we probed, H301 is known to form key hydrogen bonding with the ligand.[4,50] Interestingly, some of the residue pairs associated with L347, G348, and M351, located at the entrance of binding cleft, have to expand during the ligand binding to accommodate the ligand to the active site (Figure 5A, blue panel), reminiscent of partial unfolding and refolding observed in protein−protein or protein−ligand bindings.[12,51,52]

Intersubunit Dynamics Probed with the Residue Pairs at the Subunit Interface. The dynamics of intersubunit residue pairs is responsible for the concerted motion of tetrameric SAHH. The intersubunit residue pairs, including K166−G418, F189−Q251 at AB (or CD) interface and W17−I321, W17−K322, K20−I321 at AC (or BD) interface, exhibit a single exponential decay slower than that of the intrasubunit residue pairs (the magenta and cyan panels of Figure 5A). It is of particular note that many of individual time traces of intersubunit residue pairs show a bimodal hopping transition (Figure 5B); the histogram of $d_{AC(BD):W17-I321}$ can be described using a double Gaussian function, centered at 8 and 14 Å in the absence of ligand. Upon ligand binding, the population shifts toward $d_{AC(BD):W17-I321} \approx 8$ Å, reminiscent of population shift mechanism.

Two-Step Mechanism of Open-to-Closed Transition Dynamics. Comparison of the transition times (τ) of various residue pairs revealed an order of events that occur during the O→C dynamics. The hierarchical order of τ (Figure 5) suggests that transition dynamics progresses from intrasubunit closure dynamics to the intersubunit contact formations. Upon the ligand binding, (i) the intrasubunit residue pairs in the cleft regions are formed on time scales of $\tau \approx 33-38$ μs. Many residue pairs, particularly located at the entrance of the binding cleft (blue spheres in Figure 5A), show nonmonotonic kinetics with transition times of $\tau \approx 33-36$ μs, followed by α angle transition in 37 μs. (ii) The intersubunit residue pairs (cyan and magenta spheres in Figure 5A) and the intrasubunit residue pairs at the central region of binding cleft (green spheres) decrease with decay times of $\tau \approx 35-38$ μs. Of particular note is that the contact formation of residue pairs at the AC (or BD) interface precedes the full ordering of the binding cleft. Finally, the dimer−dimer rotation between AB and CD ($\beta \approx 4°\rightarrow 11°$) is completed in 43 μs.

To summarize (see Figures 5 and 6), in response to the ligand binding, the closure dynamics of the binding cleft due to the formation of intrasubunit contacts (blue), probed with $\langle\alpha\rangle$, precedes both the intersubunit residue contacts between AC or BD interface (cyan) and the more precise positioning of the active site (green). The rearrangement of interfacial contacts

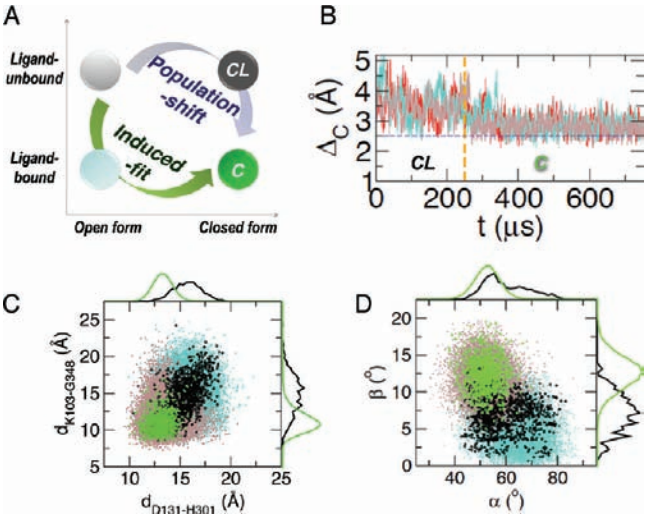


Figure 7. Comparison between the closed-like (CL) and closed structures (C). (A) A schematic diagram of molecular recognition processes contrasting induced fit and conformational selection mechanisms. (B) Exemplary time traces of Δ_C visiting the CL state prior to the ligand binding ($t < 250$ μs, $\Delta_C < R_c = 2.5$ Å) and the C state in the presence of ligand ($t > 250$ μs, $\Delta_C < R_c = 2.5$ Å). (C) Scattered plot using two intrasubunit residue distances. Ensembles of CL and C states are displayed in black and green, respectively, and their histograms are shown in the upper and right side of the graphs. For comparison, the projected space sampled by the whole trajectories before and after the ligand binding are shown in cyan and brown. (D) Scattered plot of CL and C ensembles using α and β.

(magenta) between A and B (or C and D) subunits and dimer−dimer rotation, measured with $\langle\beta\rangle$, occurs at the stage later than the closure dynamics measured with α. From local compaction at the binding cleft to the global compaction of entire structure exists a hierarchical order in O→C transition dynamics.

Transition Dynamics of SAHH in Response to the Ligand Binding Occurs via Both Modulation of Energy Landscape and Selection of Conformational Ensemble. Two limiting mechanisms, induced fit and population shift (or conformational selection) (Figure 7A), are often contrasted in literatures to account for the allosteric dynamics associated with ligand binding or protein−protein association.[53,54] In the SAHH dynamics, it is clear that the interaction between the ligand and SAHH ($H^{SAHH-lig}$) is the main cause of the O→C transition. However, it is not clear whether the ligand binding selects a small subpopulation of the closed structure from a pre-existing ensemble or modulates the folding landscapes. From the perspective of statistical thermodynamics, this question may be answered by projecting the population of accessible conformational states on a map in reduced dimension and comparing the two populations obtained from the holo and ligand-bound conditions. If the overlap between the statistical ensembles corresponding to the closed structure under the energy Hamiltonians of the holo and ligand-bound states is substantial, then population shift is more plausible. In contrast, if there is little overlap between the statistical ensembles, then induced fit mechanism is favored.

Because of the attractive contact pairs incorporated from the closed structure (see Methods Section), SAHH structure in the open form can transiently visit compact (or closed-like, CL) structures even in the absence of ligand (Figure 7B).[55] We assessed the similarity between the ensemble of CL structures

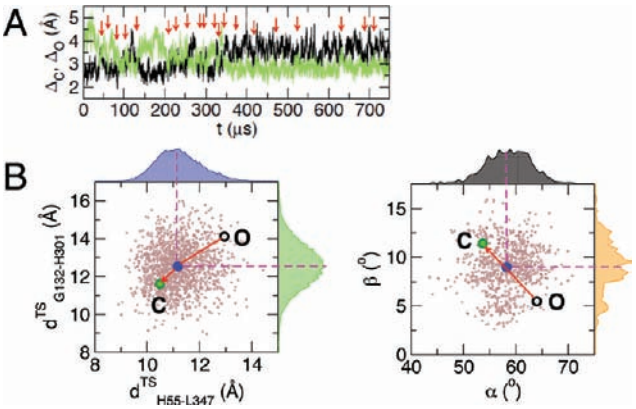


Figure 8. TSEs of the O→C transitions. (A) Time-dependent changes of Δ_O (black) and Δ_C (green) in a sample trajectory. The instances of $|\Delta_O(t) - \Delta_C(t)| < \delta$, which defines the transition state, are marked with red arrows. (B) The TSEs represented with distances of two intrasubunit residue pairs (left) and with α and β angles (right). On the two-dimensional maps representing TSE, the average positions of the TSE are marked with blue spheres, and the average distances and angles for open and closed forms are marked with the black and green circles. The calculated Tanford β-like values for the two maps are 0.33 and 0.44, respectively.

that satisfies $\Delta_C < R_c$, where $R_c = 2.5$ Å in holo structure and the ensemble of closed structures (C) that also satisfies $\Delta_C < R_c$ in the presence of ligand (Figure 7C and D). Following the line of argument made above, if the enzyme recognizes the ligand solely by selecting a pre-existing conformational ensemble, a distribution of scattered plot using inter-residue distances ($d_{K103-G346}$ and $d_{D131-H301}$) or angles (α and β) as surrogate reaction coordinates will show a substantial overlap of CL ensemble with C ensemble. On the contrary, under the induced fit mechanism, we expect a difference between the distributions of CL and C. In Figure 7C and D, C and CL structural ensembles show clear distinction but with a certain degree of overlap, giving evidence for both induced fit and population shift mechanisms.

Depending on the molecule being studied, the concentration of ligand and the position of probe being attached, one of the two mechanisms becomes more dominant.[53] For SAHH, the feature of conformational selection to the closed structure is observed in the residue pairs at the intersubunit interface (see the dynamics of $d_{AC(BD):W17-I321}$ in Figure 5B, where one of the bimodal basins corresponding to the closed structure at $d_{AC(BD):W17-I321} \approx 8$ Å becomes more populated upon the ligand binding) while the modulation of energy landscape, which alters the positions of energy minima, is more dominant in the dynamics of intrasubunit residue pairs.

Characteristics of Transition-State Ensemble of SAHH during the O→C Transition. Identifying characteristics of the transition state of an enzyme along the reaction coordinate associated with its functional motion is an important task in designing an inhibitor drug that can fit to the active site. To this end, we analyzed the transition-state ensembles (TSEs) of SAHH conformations by using $\Delta_C(t)$ and $\Delta_O(t)$ as surrogate reaction coordinates.[10,37] Upon ligand binding, $\langle\Delta_C(t)\rangle$ increases and $\langle\Delta_O(t)\rangle$ decreases, and they cross at a certain moment, i.e., $\langle\Delta_O(t)\rangle = \langle\Delta_C(t)\rangle$ (Figure 4D, left). Transition states can be collected from the individual time traces by imposing a condition $|\Delta_O(t) - \Delta_C(t)| < \delta$ with $\delta = 0.01$ Å for intrasubunit and $\delta = 0.3$ Å for intersubunit dynamics (Figure 8A).

The TSE characterized by the pair of intrasubunit residue distances ($d^{TS}_{H55-L347}$, $d^{TS}_{G132-H301}$) or angles (α and β) has a broad and heterogeneous distribution (Figure 8B), implying that the transition routes are formed on a plastic energy landscape. A Tanford β-like paramter, defined using the equilibrated distances of open form (x_O) and closed form (x_C), $\beta_T = (x_{TS} - x_C)/(x_O - x_C)$, evaluates the similarity of TSE to the ligand-bound state.[37] We obtained $\beta_T = 0.33$ and 0.44 for the dynamics of both distance and angle pairs, suggesting that the TSEs resemble the ligand-bound (closed) state ($\beta_T < 0.5$).

CONCLUDING REMARKS

Allosteric regulation of protein function is encoded in the sequence space of protein family and in the native topology of protein architecture. Functional and energetic restraints imposed by the residue network give rise to the close correlation between conformational fluctuations around the native state and transition dynamics in response to the ligand binding, although the pattern and time scale associated with these dynamics differ in each molecule due to the stochastic nature of dynamic processes and plasticity of conformational space. The analysis of various inter-residue dynamics from our BD simulations revealed that upon ligand binding, the local signals at the active site of each subunit are transmitted progressively to the intersubunit contact formation that in turn invokes the precise alignment of the catalytic elements at the active site, followed by the dimer−dimer rotation. Along with the experimental studies reporting the effect of mutation on the SAHH activity,[42,44,45] our multifaceted study of SAHH dynamics suggests that the enzymatic function is not simply represented by the relationship between the active site and ligand but is the consequence of communication between amino acids nonlocally wired over the entire structure. Together with the list of evolutionarily and topologically important residues identified by the SCA and SPM (Figures 2 and 3 and Table 1), the dynamic picture of O→C transitions of SAHH shown in this study is amenable to further experimental investigations.

ASSOCIATED CONTENT

Supporting Information. Additional figures, tables, and movies of BD simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
sunchoi@ewha.ac.kr; hyeoncb@kias.re.kr

ACKNOWLEDGMENT

This work was supported in part by the grants from the National Leading Research Lab (NLRL) program (2011-0028885), Brain Research Center of the 21st Century Frontier Research Program (2011K000289), and the National Core Research Center (NCRC) program (2011-0006244) funded by the Ministry of Education, Science, and Technology (MEST) and the National Research Foundation of Korea (NRF) to S.C., the grant from World Class University (WCU) project (R31-2008-000-10010-0) to L.S.J., and NRF grant (2010-0000602) to C.H. We thank Korea Institute for Advanced Study for providing computing resources.

REFERENCES

(1) De Clercq, E. *Nat. Rev. Drug Discovery* **2002**, *1*, 13–25.
(2) Kuntz, I. *Science* **1992**, *257*, 1078–1082.
(3) Robert, E.; Bender, S. *Chem. Rev.* **1997**, *97*, 1359–1472.
(4) Lee, K. M.; Choi, W. J.; Lee, Y.; Lee, H. J.; Zhao, L. X.; Lee, H. W.; Park, J.; Kim, H.; Hwang, K.; Heo, Y.; Choi, S.; Jeong, L. *J. Med. Chem.* **2011**, *54*, 930–938.
(5) Hardy, J.; Wells, J. *Curr. Opin. Struct. Biol.* **2004**, *14*, 706–715.
(6) Goodey, N.; Benkovic, S. *Nat. Chem. Biol.* **2008**, *4*, 474–482.
(7) Swain, J.; Gierasch, L. *Curr. Opin. Struct. Biol.* **2006**, *16*, 102–108.
(8) Knowles, J. *Nature* **1991**, *350*, 121–124.
(9) Schnell, J.; Dyson, H.; Wright, P. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 119–140.
(10) Chen, J.; Dima, R. I.; Thirumalai, D. *J. Mol. Biol.* **2007**, *374*, 250–266.
(11) Johnson, D. A.; Akamine, P.; Radzio-Andzeim, E.; Madhusudan; Taylor, S. S. *Chem. Rev.* **2001**, *101*, 2243–2270.
(12) Hyeon, C.; Jennings, P. A.; Adams, J. A.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3023–3028.
(13) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, J.; Karplus, M.; Kern, D. *Nature* **2007**, *450*, 913–916.
(14) Whitford, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N. J. *Mol. Biol.* **2007**, *366*, 1661–1671.
(15) Kamata, K.; Mitsuya, M.; Nishimura, T.; Eiki, J.; Nagata, Y. *Structure* **2004**, *12*, 429–438.
(16) Frauenfelder, H.; McMahon, B.; Austin, R.; Chu, K.; Groves, J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2370–2374.
(17) Osborne, M.; Schnell, J.; Benkovic, S.; Dyson, H.; Wright, P. *Biochemistry* **2001**, *40*, 9846–9859.
(18) Henzler-Wildman, K.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M.; Petsko, G.; Karplus, M.; Hübner, C.; Kern, D. *Nature* **2007**, *450*, 838–844.
(19) Hammes-Schiffer, S.; Benkovic, S. *Annu. Rev. Biochem.* **2006**, *75*, 519–541.
(20) Bhabha, G.; Lee, J.; Ekiert, D.; Gam, J.; Wilson, I.; Dyson, H.; Benkovic, S.; Wright, P. *Science* **2011**, *332*, 234–238.
(21) Lockless, S. W.; Ranganathan, R. *Science* **1999**, *286*, 295–299.
(22) Halabi, N.; Rivoire, O.; Leibler, S.; Ranganathan, R. *Cell* **2009**, *138*, 774–786.
(23) Zheng, W.; Brooks, B. R.; Doniach, S.; Thirumalai, D. *Structure* **2005**, *13*, 565–577.
(24) Balabin, I.; Yang, W.; Beratan, D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14253–14258.
(25) Dhillon, A.; Meikle, S.; Yazici, Z.; Eulitz, M.; Kolch, W. *EMBO J.* **2002**, *21*, 64–71.
(26) Hanson, J. A.; Duderstadt, K.; Watkins, L. P.; Bhattacharyya, S.; Brokaw, J.; Chu, J.-W.; Yang, H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18055–18060.
(27) Lee, J.; Natarajan, M.; Nashine, V.; Socolich, M.; Vo, T.; Russ, W.; Benkovic, S.; Ranganathan, R. *Science* **2008**, *322*, 438.
(28) Hyeon, C.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2175–2180.
(29) Hyeon, C.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17382–17387.
(30) Chen, J.; Darst, S. A.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 12523–12528.
(31) Süel, G.; Lockless, S.; Wall, M.; Ranganathan, R. *Nat. Struct. Mol. Biol.* **2002**, *10*, 59–69.
(32) Dima, R. I.; Thirumalai, D. *Protein Sci.* **2006**, *15*, 258–268.
(33) Porter, D. *J. Biol. Chem.* **1993**, *268*, 66.
(34) Yang, X.; Hu, Y.; Yin, D.; Turner, M.; Mengmeng Wang, O.; Borchardt, R.; Howell, P.; Kuczera, K.; Schowen, R. *Biochemistry* **2003**, *42*, 1900–1909.
(35) Hyeon, C.; Thirumalai, D. *Nat. Commun.* **2011**, *2*, 487.
(36) Hyeon, C.; Dima, R. I.; Thirumalai, D. *Structure* **2006**, *14*, 1633–1645.
(37) Hyeon, C.; Lorimer, G. H.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18939–18944.
(38) Edgar, R. *BMC bioinformatics* **2004**, *5*, 113.
(39) Getz, G.; Levine, E.; Domany, E. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12079.
(40) Getz, G.; Domany, E. *Bioinformatics* **2003**, *19*, 1153–1154.
(41) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352–1369.
(42) Yuan, C.; Ault-Riché, D.; Borchardt, R. *J. Biol. Chem.* **1996**, *271*, 28009–28016.
(43) Li, Q.; Cai, S.; Borchardt, R.; Fang, J.; Kuczera, K.; Middaugh, C.; Schowen, R. *Biochemistry* **2007**, *46*, 5798–5809.
(44) Ault-Riché, D.; Yuan, C.; Borchardt, R. *J. Biol. Chem.* **1994**, *269*, 31472–31478.
(45) Vugrek, O.; Belužić, R.; Nakić, N.; Mudd, S. *Hum. Mutat.* **2009**, *30*, E555–E565.
(46) Wang, M.; Borchardt, R.; Schowen, R.; Kuczera, K. *Biochemistry* **2005**, *44*, 7228–7239.
(47) Hu, C.; Fang, J.; Borchardt, R.; Schowen, R.; Kuczera, K. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 131–143.
(48) Wang, M.; Unruh, J.; Johnson, C.; Kuczera, K.; Schowen, R.; Borchardt, R. *Biochemistry* **2006**, *45*, 7778–7786.
(49) Yin, D.; Yang, X.; Hu, Y.; Kuczera, K.; Schowen, R.; Borchardt, R.; Squier, T. *Biochemistry* **2000**, *39*, 9811–9818.
(50) Wang, T.; Lee, H.; Tosh, D.; Kim, H.; Pal, S.; Choi, S.; Lee, Y.; Moon, H.; Zhao, L.; Lee, K.; Jeong, L. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4456–4459.
(51) Miyashita, O.; Onuchic, J. N.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12570–12575.
(52) Schrank, T.; Bolen, D.; Hilser, V. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16984–16989.
(53) Hammes, G. G.; Chang, Y. C.; Oas, T. G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13737–13741.
(54) Itoh, K.; Sasai, M. *Proc. Natl. Acad. Sci.* **2010**, *107*, 7775–7780.
(55) Bahar, I.; Chennubhotla, C.; Tobi, D. *Curr. Opin. Struct. Biol.* **2007**, *17*, 633–640.