Fly-casting with ligand–sliding and orientational selection to support the complex formation of a GPCR and a middle-sized flexible molecule

Junichi Higo^{1,2,*}, Kota Kasahara³, Gert-Jan Bekker⁴, Benson Ma⁴, Shun Sakuraba⁵, Shinji Iida⁶, Narutoshi Kamiya¹, Ikuo Fukuda¹, Hidetoshi Kono⁵, Yoshifumi Fukunishi⁷, Haruki Nakamura⁴

¹ Graduate School of Information Science, University of Hyogo, 7-1-28 Minatojima Minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

² Research Organization of Science and Technology, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577, Japan

³ College of Life Sciences, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577, Japan

⁴ Institute for Protein Research, Osaka University, 3-2 Yamada-oka, Suita, Osaka 565-0871, Japan

⁵ Institute for Quantum Life Science, National Institutes for Quantum Science and Technology, 8-1-7 Umemi-dai, Kizu, Kyoto 619-0215, Japan

⁶ Technology Research Association for Next-Generation Natural Products Chemistry, 2-3-26, Aomi, Koto-ku, Tokyo, Japan 135-0064

⁷ Cellular and Molecular Biotechnology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7, Aomi, Koto-ku, Tokyo, 135-0064, Japan

*Correspondence and requests for materials should be addressed to J.H. (e-mail: higo@protein.osaka-u.ac.jp)

Abstract

To elucidate computationally a binding mechanism of a middle-sized flexible molecule, bosentan, to a GPCR protein, human endothelin receptor type B (hETB), a GA-guided multidimensional virtual-system coupled molecular dynamics (GA-mD-VcMD) simulation was performed. This method is one of generalized ensemble methods and produces a free-energy landscape of the ligand-receptor binding by searching large-scale motions accompanied with stably keeping the fragile cell-membrane structure. All molecular components (bosentan, hETB, membrane, and solvent) were represented with an all-atom model, and sampling was carried out from conformations where bosentan was distant from the binding site in the hETB's binding pocket. The deepest basin in the resultant free-energy landscape was assigned to the native-like complex conformation. The obtained binding mechanism is as follows. First, bosentan fluctuating randomly in solution is captured by a tip region of the flexible N-terminal tail of hETB via nonspecific attractive interactions (fly-casting). Bosentan then occasionally slides from the tip to root of the N-terminal tail (ligand-sliding). In this sliding, bosentan passes the gate of the binding pocket from outside to inside of the pocket with accompanying a quick reduction of the molecular orientational variety of bosentan (orientational selection). Last, in the pocket, ligand-receptor attractive native contacts are formed, and eventually the native-like complex is completed. The bosentan-captured conformations by the tip- and root-regions of the N-terminal tail correspond to two basins in the free-energy landscape, and the ligand-sliding corresponds to overcoming a free-energy barrier between the basins.

Introduction

G protein-coupled receptors (GPCRs) are membrane proteins constructing a large evolutionarily related protein family with a variety of molecular functions¹. In general, GPCR has seven transmembrane helices (TMH1 to TMH7), which are embedded in membrane. GPCRs activate cellular responses via detecting molecules outside the cell. Because GPCRs are related to many diseases, they have been important targets for drugs². Endothelin-1 (ET1) is a peptide (21 residues long) with a strong vasoconstrictor action discovered in humans³. For exerting the activity, ET1 transmits signals by interacting with two GPCRs: The endothelin type A (ETA)⁴ and endothelin type B (ETB)⁵ receptors. The complex structure of human endothelin type-B receptor (hETB) and ET1 was solved using X-ray crystallography⁶. In this complex, ET1 is bound to a binding pocket of hETB on cytoplasmic side.

Bosentan competes with ET1 when binding to human ETA (hETA) and human hETB (hETB), and inhibits the ET1's vasoconstriction effect as an antagonist^{7,8,9}. The

tertiary structure of the bosentan-hETB complex was solved by X-ray crystallography¹⁰. In this structure, bosentan binds to the binding pocket of hETB more deeply than ET1. However, the interaction patterns between the C-terminal segment of ET1 and amino-acid residues in the binding pocket of hETB are similar with those in the bosentan-hETB complex.

Bosentan is a medium-sized (551.6 Da) drug molecule and relatively large comparing to other commercial drug molecules. This molecule is flexible because it has long flexible sidechains (figure 1a) whereas the central ring, to which the sidechains are connected, is stiff. Thus, bosentan will adopt temporal complex conformations (encounter complexes) before reaching the deep position of the hETB's pocket (i.e., the native-complex position). To investigate those temporal conformations, a special molecular dynamics (MD) simulation is suitable that can sample various conformations in the complex-formation process.

Recently, we introduced a generalized ensemble method, a multi-dimensional virtual-system coupled molecular dynamics (mD-VcMD) simulation¹¹, which enhances conformational sampling with repetition of simulations (iteration) in a multi-dimensional (mD) reaction-coordinate (RC) space. Repeating the iterations, the sampled volume in the mD-RC space is expanded and sampled conformations (snapshots) are collected from the iterations. Importantly, mD-VcMD assigns a statistical weight (thermodynamic weight equilibrated at a simulation temperature) to each of the snapshots. Therefore, the ensemble of snapshots is equilibrated, and any equilibrated physicochemical quantity is computable from the ensemble if the quantity is expressed by the system's coordinates. We call this simulation method the “original mD-VcMD method” or simply the “original method” in this paper. The original method was applied to a large and complicated system consisting of ET1 and hETB, where hETB was embedded in an explicit membrane and the membrane was immersed in an explicit solvent¹².

Although the original mD-VcMD was applicable to the large and complicated system, we encountered a difficulty: Once a poorly sampled region emerged in the mD-RC space in an iterative simulation, the conformation might be trapped in this region in the subsequent iteration. Thus, a more robust method was required to proceed the sampling. To overcome this difficulty, we introduced a GA-guided multi-dimensional virtual-system coupled molecular dynamics method (GA-mD-VcMD)¹³, which is an extension of the original mD-VcMD, where “GA” is an abbreviation of genetic algorithm. We applied this method to some biological systems, which consist of a protein and ligand in an explicit solvent^{14,15,16}.

From the original mD-VcMD applied to binding of ET1 to hETB, two binding mechanisms were found: Fly casting^{17,18,19} and conformational selection²⁰. The receptor hETB has an intrinsically disordered N-terminal tail. The floating ET1 in solution was captured by the fluctuating N-terminal tail (fly-casting mechanism) because the interactions between ET1 and the hETB's N-terminal were attractive¹². We note that the root of the N-terminal tail is located at the gate of the binding pocket of hETB. This indicates that once ET1 is captured by the N-terminal tail, the ligand is restricted in a narrow volume around the gate of the binding pocket, resulting in increase of the possibility that ET1 contacts the gate of the binding pocket. Next, only a fraction of various orientations of ET1 were allowed to enter the hETB pocket (conformational selection). This suggests that the pocket size is not large enough for ET1 to rearrange the molecular orientation in the pocket, and that only ET1 with molecular orientations advantageous for the native complex formation is accepted.

In the present study, we sampled the interactions between the medium-sized molecule bosentan and hETB by the GA-mD-VcMD simulation. Except for the difference of ligand and simulation method, there are two differences between the previous and present studies. (i) We prolonged the intrinsically disordered N-terminal tail of hETB of the X-ray structure^{6,10}: Five and ten residues long for the previous and the current studies, respectively¹². This tail extension varies the feature of the fly-casting mechanism as shown later. Because GPCRs have a long N-terminal tail in general²¹, this extension of the tail may support experimental study on the role of the N-terminal tail of GPCRs^{22,23}. Second, the initial simulation conformations for the present study were those where bosentan was distant from hETB while those for the previous study were the native-like complex conformations of ET1 and hETB. In general, the difference of initial conformations significantly influences the resultant conformational ensemble. The current study is not guaranteed if the native-like conformations are sampled. The native-like complex should be *discovered* as the lowest free-energy conformation out of many possible ones. Contrarily, in the previous study, the native-like complex is always involved in the resultant ensemble, which means that the native-like complex may not necessarily correspond to the lowest free-energy conformation because a lower free-energy conformation may be overlooked in the sampling.

Our GA-mD-VcMD simulation produced a free-energy landscape of the bosentan-hETB system, and probability distribution functions of some quantities were calculated from the resultant conformational ensemble. Because the free-energy landscape covered both unbound conformations and the native-like complex structures, and because the lowest free-energy conformation corresponded to the native-like

complex structure, we could analyze entirely the complex formation. We discuss the generality of the binding mechanism found in bosentan-hETB to the ligand-GPCR binding. Because bosentan is a drug molecule, the present study helps drug development research not only by proposing the lowest free-energy conformations but also by investigating the binding process.

Methods

The aim of this work is to investigate the complex-formation process of bosentan and hETB by the GA-mD-VcMD method. First, we briefly explain this method, which provides a conformational ensemble of the system and a thermodynamic weight at the simulation temperature T (300 K in the present study), which is assigned to conformations stored in the ensemble. Details of GA-mD-VcMD have been explained elsewhere^{13,14,15,16}. Next, we explain the molecular system which consists of hETB, bosentan, membrane (cholesterol and POPC lipid molecules), and solvent (water molecules and ions). Then, we describe multiple reaction coordinates introduced and the initial conformations for simulation. Lastly, we account for the method to calculate the spatial density of bosentan around hETB.

GA-mD-VcMD and a conformational ensemble

To sample extensively both the unbound and bound conformations of the system, GA-mD-VcMD controls the system's motions in a multi-dimensional (mD) space constructed by multiple reaction coordinates (RCs). Definition of an RC is presented in Section 1 of Supplementary Information (SI) and figure S1. We introduce three RCs denoted as $\lambda^{(h)}$ ($h = \alpha, \beta$, or γ), whereas the dimensionality of the RC space is arbitrary in theory¹³. Therefore, the word “multi-dimensional (mD)” indicates three-dimensional (3D) in the present study. The actual definition for the three RCs is given later.

Conformational sampling of the system by GA-mD-VcMD at 300 K produces a distribution (density) of the system in the mD-RC space (not in the real space). Because the conformational space of a biological molecular system is wide, the GA-mD-VcMD simulation is executed iteratively, by which the sampled volume of 3D-RC space is expanded. Here, the density at a 3D-RC position $[\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)}]$ is denoted as $Q_{cano}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$, and the density computed from the M -th iteration may be expressed as $Q_{cano}^{(M)}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$ when necessary.

The simulation is terminated when a convergence criterion is satisfied for $Q_{cano}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$. Snapshots of the system are saved from all iterations. After the convergence, a thermodynamic weight at the simulation temperature T is assigned to

each saved snapshot based on $Q_{cano}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$. Therefore, the snapshots construct a thermally equilibrated conformational ensemble (canonical ensemble) at T (300 K).

Molecular system studied

We generated the molecular system consisting of bosentan, hETB, cholesterol surrounding hETB, POPC bilayer where hETB is embedded, and an explicit solvent (water molecules, Na^+ and Cl^- ions). All of them were packed in a periodic boundary box. Details for the system preparation are given in Section 2 of SI. figure S2 is a diagram outlining the generation of the molecular system.

Two reference crystal structures are available for generating the molecular system for the simulation: The bosentan–hETB (PDB ID 5xpr; 3.6 Å resolution) and K8794–hETB (5x93; 2.2 Å) complex structures. Importantly, the two crystal structures are considerably similar each other (figure S3a) because K8794 is an analog of bosentan (figure S3b). Because K8794–hETB complex has a better resolution than bosentan–hETB complex, we used the K8794–hETB complex to prepare for the molecular system.

First, we replaced K8794 by bosentan. This complex conformation is referred to “Complex structure of bosentan–hETB” in figure S2. In the crystallography, some mutations were applied to the amino-acid sequence of hETB to increase the quality of the diffraction data¹⁰. Then, we returned the hETB’s sequence to the wild type, and embedded the complex in the POPC bilayer. Four cholesterol were introduced into the hETB–membrane interface (see Section 2 of SI). Then, the system was immersed in a periodic boundary box filled by water molecules. The x-, y-, and z-coordinate axes for specifying the atomic positions were set parallel to the sides of the periodic box (the x-y plain was parallel to the membrane surface). The box dimension was 71.33 Å (x-axis), 71.33 Å (y-axis), and 132.17 Å (z-axis). Finally, sodium and chlorine ions were introduced by replacing water molecules randomly with the ions to neutralize the system at a physiological ion concentration of 130 mM. This conformation was named “Complex structure in periodic box” in figure S2. The total number of atoms of the system was 69,062: numbers of protein and bosentan atoms were 5,289 and 68, respectively; four cholesterol (296 atoms); 127 POPC molecules (17,018 atoms); 15,440 water molecules (46,320 atoms); and 28 sodium and 43 chlorine ions.

After energy minimization of the above system, followed by NVT and NPT simulations, we obtained the complex structure shown in figure S4, whose resultant periodic-box size was 69.37 Å × 69.37 Å × 140.29 Å. This is represented as “Complex structure in figure S4 (native-complex structure)” in figure S2. Because this complex structure is a relaxed conformation of the native complex, bosentan was located at a

position close to that in the X-ray structures (PDB ID: 5xpr and 5x93). Therefore, we call this conformation as the “native complex structure” in the present study. It should be noted that this native complex structure is *not* the initial conformation of the current GA-mD-VcMD simulations. As explained later, we generated conformations where bosentan was completely separated from the binding pocket of hETB. They were the conformations that we used as the initial simulation conformations of GA-mD-VcMD.

Set of Three RCs Introduced to Control System's Motions

Before describing the generation of the initial conformations of GA-mD-VcMD, we explain concretely the set of three RCs ($\lambda^{(h)}$; $h \in \{\alpha, \beta, \gamma\}$) introduced for the current study. Setup of RCs is essential to increase the sampling efficiency though RCs can be set arbitrary in theory. We imposed three conditions on the RCs: The RCs can control (i) the gate opening/closing of the hETB's binding pocket, (ii) the ligand approaching/departing to the receptor, and (iii) the ligand orientation relative to the receptor.

Although various RCs are possible to satisfy the three conditions above, we set them in a straightforward manner as shown in figures 1b and c, and the exact specification for the atom groups are given in the figure caption. We briefly explain the roles of the RCs. For the condition (i), two atom groups G_α^A and G_α^B (figure 1b) are used to define the first RC $\lambda^{(\alpha)}$ that describes open /close of the gate of the hETB's binding pocket. figure 1c shows atom groups G_β^A and G_β^B to define $\lambda^{(\beta)}$, as well as G_γ^A and G_γ^B to define $\lambda^{(\gamma)}$. The groups G_β^A and G_β^B are close to each other in the native complex, and G_γ^A and G_γ^B are as well. Contrarily, those atom groups are distant from each other in an unbound conformation. Therefore, the variations of $\lambda^{(\beta)}$ and $\lambda^{(\gamma)}$ control bosentan approaching/departing to hETB: The condition (ii) is satisfied. Lastly, the condition (iii) is satisfied when $\lambda^{(\beta)}$ increases/decreases with decreasing/increasing $\lambda^{(\gamma)}$. Figure 1a is closeup of bosentan to clarify G_β^B and G_γ^B .

Fig. 1. (A) Chemical structure of bosentan. Two atom groups G_{β}^B and G_{γ}^B , used to set reaction coordinates (RCs) $\lambda^{(\beta)}$ and $\lambda^{(\gamma)}$, are shown. (B) The first RC, $\lambda^{(\alpha)}$, defined by distance between centroids of atom groups G_{α}^A (red-colored segment in the N-terminal tail of hETB surrounded by dotted line; residues 85–89) and G_{α}^B (magenta-colored segment in β -hairpin of hETB surrounded by another dotted line; residues 243–254). Binding pocket of hETB is indicated by blue-colored broken-line circle. Bound bosentan is shown at the bottom of the pocket. Green and cyan regions are defined in panel (C). (C) The second and third RCs, $\lambda^{(\beta)}$ and $\lambda^{(\gamma)}$, respectively. $\lambda^{(\beta)}$ is defined by distance between centroids of G_{β}^A (a cyan-colored part of the fifth transmembrane helix TMH5; residues 273–281) and G_{β}^B (a part of bosentan surrounded by dotted line). $\lambda^{(\gamma)}$ is defined by G_{γ}^A (a green-colored part of the seventh transmembrane helix TMH7; residues 372–379) and G_{γ}^B (other part of bosentan surrounded by the other dotted line). Bound and unbound bosentans are displayed.

Initial conformations for GA-guided mD-VcMD

We explain generation of the initial conformations of GA-mD-VcMD. Our aim is to explore the conformational space from the unbound to bound states. Thus, the initial conformations should not be the native complex structure but be a completely unbound conformations. Since we had only the native complex structure (figure S4) at this stage, we needed to generate the unbound conformations from the complex structure. For this purpose, we performed ten simulations (called “separation simulations”) starting from the native complex structure with imposing three requirements to RCs to the system’s conformation: $24.0 \text{ \AA} \leq \lambda^{(\alpha)} \leq 25.0 \text{ \AA}$, $60.0 \text{ \AA} \leq \lambda^{(\beta)} \leq 65.0 \text{ \AA}$, and $60.0 \text{ \AA} \leq \lambda^{(\gamma)} \leq 65.0 \text{ \AA}$. The three RCs for the native complex structure are $\lambda^{(\alpha)} = 8.22 \text{ \AA}$, $\lambda^{(\beta)} = 8.35 \text{ \AA}$, and $\lambda^{(\gamma)} = 6.21 \text{ \AA}$, respectively. During the separation simulations, bosentan moved from the binding pocket to unbound conformations. The resultant conformations are shown in figure S5, where bosentans are completely separated from hETB, which are represented “Unbound conformations in figure S5” in figure S2.

Next, for further relaxing those unbound conformations, we performed 2,000 runs in total: 200 (1×10^6 steps at 300 K) runs starting from each of those ten unbound conformations with different random seeds to the initial atomic velocities. In the 2,000 runs, we applied the following three restrictions to RCs: $\lambda^{(\alpha)} \leq 25.0 \text{ \AA}$, $\lambda^{(\beta)} \leq 65.0 \text{ \AA}$, and $\lambda^{(\gamma)} \leq 65.0 \text{ \AA}$. The upper limits were introduced to prevent bosentan from flying further away from hETB in the periodic box. We refer to those 2,000 runs as “diffusion simulations”. The last snapshots from 2,000 runs were used as the initial conformations of the GA-mD-VcMD simulations. Figure S6 plots the positions of the last snapshots of the simulations in the 3D-RC space and the native complex structure in the 3D-RC space.

Those conformations are represented as “Distributed conformations in figure S6” in figure S2.

It should be noted that figure S6 demonstrates that the conformations sampled in the diffusion simulations did not reach the native complex structure. We emphasize that all the current results were not influenced by the structural features of the native complex because the 2,000 initial conformations are those diffused freely from the ten completely unbound conformations (figure S5).

GA-mD-VcMD simulations

As mentioned in the INTRODUCTION–section, GA-mD-VcMD enhances conformational sampling in the 3D-RC space, and the thermodynamic weights of sampled conformations at the simulation temperature (300 K) are computed using the converged distribution $Q_{cano}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$.

In the original mD-VcMD¹² and GA-mD-VcMD¹³, each RC axis $\lambda^{(h)}$ ($h \in \{\alpha, \beta, \gamma\}$) is divided into small 1D zones (the zone was named originally as “virtual zone”¹³), and accordingly the 3D-RC space is divided into small 3D zones. Table S1 lists the actual division of each 1D axis into zones. During a short time–interval $\Delta\tau_{trns}$ of simulation, the system’s conformation freely moves only in a single 3D zone with a restoring force outside the zone. Then, an inter-zone transition from the zone to one of neighboring zones is attempted every time interval. In the present study, we set $\Delta\tau_{trns} = 20$ ps, which corresponds to 1×10^4 simulation steps with time step of 2 fs. The inter-zone transition probabilities are determined using Q_{cano} ^{12,13}. In this way, the conformation moves in a wide 3D-RC space by the inter-zone transitions.

We performed N_{run} ($= 2,000$) runs at 300 K in parallel starting from different conformations in each iteration to further increase the sampling efficiency. A simple integration of the N_{run} trajectories can be regarded as a single, long trajectory by connecting the trajectories in arbitrary order^{24,25}. The N_{run} initial conformations for the first iteration of GA-mD-vcMD are the last conformations from the diffusion simulations (figure S6). We obtain snapshots from the first to M -th iterations when the M -th iteration has finished. The N_{run} initial conformations for the $(M + 1)$ -th iteration are selected from those snapshots so that they distribute as evenly as possible in the 3D-RC space. On the other hand, when a certain RC region is poorly sampled, we pick more snapshots from or near the poorly sampled RC region than from well-sampled RC regions¹³.

We set the simulation length of each run to 1×10^6 steps ($1 \times 10^6 \times 2$ fs = 2.0 ns). Thus, the total length for an iteration is 4.0 μ s ($= 2.0$ ns $\times 2,000$). We performed 55 iterations in the present study, which corresponds to 220.0 μ s

($4.0 \mu\text{s} \times 55$) in total. A snapshot was stored every 1×10^5 steps (0.2 ns) in each run, and then 1.1×10^6 (220.0 μs /0.2 ns) snapshots were stored in total. We calculated distribution functions of various quantities at the simulation temperature (300 K) from the resultant ensemble of snapshots.

The GA-mD-VcMD algorithm¹³, which was initially available in a MD simulation program omegagene/myPresto²⁶, has now been implemented in GROMACS²⁷ using the PLUMED plug-in for the current study²⁸. The simulation condition is as follows: LINCS²⁹ to constrain the covalent-bond lengths related to hydrogen atoms, the Nosé–Hoover thermostat^{30,31} to control the simulation temperature, the zero-dipole summation method^{32,33,34} to compute the long-range electrostatic interactions accurately and quickly, a time-step of 2 fs ($\Delta t = 2$ fs), and simulation temperature of 300 K. All simulations were performed on the TSUBAME3.0 supercomputer at the Tokyo Institute of Technology with using GP-GPU.

The force fields used are AMBER99SB-ILDN³⁵ for hETB and AMBER LIPID14³⁶ for the cholesterol and POPC lipid molecules, TIP3P³⁷ for water molecules and Joung–Cheatham model³⁸ for ions (chloride and sodium), respectively. The bosentan's force field was prepared as follows: First, the bosentan's atomic partial charges were calculated quantum-chemically using Gaussian09³⁹ at the HF/6-31G* level, followed by RESP fitting⁴⁰. Then, the obtained partial charges were incorporated into a GAFF (general amber force field) force-field file⁴¹, which was designed to be compatible with the conventional AMBER force-fields. The chemical structure of bosentan is shown in figure 1A.

Spatial density of bosentan around hETB

GA-mD-VcMD simulation produces a canonical ensemble of conformations each of which a thermodynamic weight at 300 K is assigned to. Therefore, using the ensemble, we can calculate a spatial density of bosentan's centroid around hETB at 300 K using the ensemble. The detailed method to calculate the spatial density is presented in Subsection 3.1 of SI.

We explain briefly notations used to express the spatial density. The 3D real space (not the 3D-RC space) is divided into cubes (size is $3 \text{ \AA} \times 3 \text{ \AA} \times 3 \text{ \AA}$), whose centers are denoted as $\mathbf{r}_{cube}$. The spatial density $\rho_{CMB}(\mathbf{r}_{cube})$ is the probability of detecting the bosentan's centroid $\mathbf{r}_{CMB}$ in the cube centered at $\mathbf{r}_{cube}$.

Results and Discussion

We repeated 55 iterations of GA-mD-VcMD with saving 1.1×10^6 snapshots from the first to 55-th iterations in total (see Methods section). First, we show that the GA-mD-VcMD produced a distribution of the system that covered both the unbound and native-like complex conformations. We next show that the highest density spot (the lowest free-energy spot) corresponds to the native-like complex structure. We then analyzed the binding process from the unbound to intermediates, and from the intermediates to the native-like complex. For this purpose, the binding process is divided into three stages: One is the early stage of binding in that bosentan is outside the binding pocket. The second is the stage where bosentan is passing the gate of the binding pocket. The last is the stage where bosentan is moving in the binding pocket.

Conformational distribution in 3D-RC space

GA-mD-VcMD produced the distribution $Q_{cano}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$ equilibrated at 300 K in the 3D-RC space (figure S8). Importantly, the distribution covered both the unbound and bound conformations, and the native-complex structure was located at the periphery of the lowest free-energy basin. Section 4 of SI explains the features of this distribution.

The method to check the convergence of the distribution is presented in Section 5 of SI. Figure S9 shows that iterations 1–10 (figures S9a–c) sampled a large volume of the 3D-RC space involving the unbound and bound conformations. Therefore, the simulation could be quitted at the tenth iteration. However, to obtain snapshots sufficient for analysis, we continued the simulation up to the 55-th iteration. Remember that we stored a snapshot every 0.2 ns (100,000 steps) of simulation.

Next, we compared the distribution $Q_{cano}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$ with that obtained from the previous study of the ET1–hETB system¹². See Section 6 of SI for details of the comparison method. Figure S10 shows that the free-energy slope along the $\lambda^{(\alpha)}$ axis (the motion along the gate opening/closing motion of the hETB's binding pocket) is gentler than those along the $\lambda^{(\beta)}$ or $\lambda^{(\gamma)}$ axis (the bosentan's approaching/departing motions to hETB). This result agrees with that from the previous study. However, the slope in the current bosentan–hETB landscape was about one-tenth of that in the previous ET1–hETB landscape. We discuss later this inconsistency between the two studies.

Three-dimensional (3D) spatial distribution of bosentan around hETB.

Whereas GA-mD-VcMD samples the mD-RC space, the conformational distribution in the mD-RC space does not always provide a concrete or intuitive image for the molecular conformation of the system. Therefore, we converted the distribution $Q_{cano}(\lambda^{(\alpha)}, \lambda^{(\beta)}, \lambda^{(\gamma)})$ (figure S8) to the conformational distribution $\rho_{CMB}(\mathbf{r}_{cube})$

defined in the 3D real space using a thermodynamic weight assigned to all the sampled snapshots using equation 31 of Ref. 13. See Subsection 3.1 of SI for the method to calculate $\rho_{CMb}(\mathbf{r}_{cube})$. Figure 2 illustrates the spatial patterns of $\rho_{CMb}(\mathbf{r}_{cube})$ of the bosentan's centroid $\mathbf{r}_{CMb}$ around hETB. Importantly, the highest-density spot (the red-colored contours; $\rho_{CMb}(\mathbf{r}_{cube}) \geq 0.5$) well agreed with the position of bosentan's centroid in the native complex structure. This result is important because the reliability of the simulation data reduces unless the highest-density spot is assigned to the native bosentan's position.

The binding pocket corresponds to the region of $\rho_{CMb}(\mathbf{r}_{cube}) \geq 0.1$ (blue-colored contours), which involved the highest-density spot. Outside of the binding pocket was characterized by a lower density of $\rho_{CMb}(\mathbf{r}_{cube}) \geq 0.01$ (magenta-colored ones), which was named “density hat” or “hat” simply. We denote the hat–pocket boundary (cyan-colored broken line in Figure 2) as “gate” of binding pocket.

Figure S11 illustrates the spatial density for the whole sampled region. Bosentan was distributed even in far regions from hETB as indicated by the low density: $\rho_{CMb}(\mathbf{r}_{cube}) \geq 0.0001$. The bosentan's density increased rapidly with approaching the binding pocket from the far region. This figure demonstrates that GA-mD-VcMD is powerful in sampling the wide conformational space efficiently.

Fig. 2. Spatial density, $\rho_{CMb}(\mathbf{r}_{cube})$, of bosentan's centroid around hETB. Iso-density map is presented at three contour levels by different colors as shown in inset. Molecular structure is the native complex (figure S4) omitting solvent. Bosentan is shown by cyan-colored stick model, and its centroid is shown by small black sphere. N-terminal tail and β -hairpin of hETB are indicated with labels. Red-colored contours are

named “highest-density spot”. Blue-colored ones ($\rho_{CMB}(\mathbf{r}_{cube}) \geq 0.1$) specify well the binding pocket of hETB. Magenta-colored region above the pocket was named “hat”. Cyan-colored broken-line indicates boundary between hat–pocket boundary, which is called “gate” of the binding pocket. See also figure S11, which illustrates $\rho_{CMB}(\mathbf{r}_{cube})$ with adding two low-density contours of $\rho_{CMB}(\mathbf{r}_{cube}) \geq 0.001$ and 0.0001.

RMSD of Bosentan between Snapshot and the Native Complex

Above we showed visually agreement of the computed highest-density spot and the bosentan’s native-complex position. Here, to analyze quantitatively the agreement, we calculated two types of root-mean-square-deviation $RMSD$ of bosentan between a snapshot and the native complex: $RMSD = RMSD_{whole}$ or $RMSD = RMSD_{core}$. $RMSD_{whole}$ is calculated using all the heavy atoms in bosentan, and $RMSD_{core}$ is done using the heavy atoms in the bosentan’s core region. Definition of the core region is given in figure 3a.

We calculated two probability distribution functions to analyze $RMSD$: Distance distribution function $P_{DDF}(RMSD)$ and radial distribution function $P_{RDF}(RMSD)$. See Subsection 3.2 of SI for details of P_{DDF} and P_{RDF} . The resultant distributions are presented in figure 3. For the whole bosentan, the highest-peak positions of P_{RDF} and P_{DDF} are at $RMSD_{whole} \approx 3.25 \text{ \AA}$ and $RMSD_{whole} \approx 5.75 \text{ \AA}$, respectively. Similarly, for the core region, those are at $RMSD_{core} \approx 2.25 \text{ \AA}$ and $RMSD_{core} \approx 2.75 \text{ \AA}$, respectively. The peak positions for the core region were smaller than those for the whole bosentan. This is because of the disorder of the longest sidechain of bosentan, as shown later.

Fig. 3. (a) Bosentan’s “core region”, which is surrounded by red-colored broken line. (b) Distance distribution function $P_{DDF}(RMSD)$ and (c) radial distribution function $P_{RDF}(RMSD)$ are presented for both of $RMSD_{whole}$ and $RMSD_{core}$. Definition of $RMSD$, P_{DDF} , and P_{RDF} is given in Subsection 3.2 of SI.

Bosentan–tail contact ratio

The structure of the N-terminal tail is not determined in the x-ray structures (PDB IDs: 5x93 and 5xpr) because the tail is intrinsically disordered. We presume that the tail may affect the binding process because the root of the tail is located at the entrance of the hETB's binding pocket. Some experiments report that an extracellular loop of GPCRs plays a role of lid to block the gate of the binding pocket^{42,43}.

Figure S7b highlights the gate of the hETB's binding pocket. This figure shows that a half of the gate, which is covered by the white solid-line rectangle, is closed by the contact of the N-terminal tail and the β -hairpin, and the other half of the gate, which is covered by the white broken-line rectangle, is open. The definition of the N-terminal tail and the β -hairpin is given in the caption of figure S7b. We can imagine readily from this figure that the gate situation can vary during the fluctuations of the tail. Below, we investigate the tail's fluctuations and the bosentan-tail contact.

First, to visualize the fluctuations of the N-terminal tail, we calculated a spatial density function $\rho_{CMNt}(\mathbf{r}_{cube})$ for the tip of the N-terminal tail. We also calculated the spatial density function $\rho_{CM\beta h}(\mathbf{r}_{cube})$ for the tip of the β -hairpin because the β -hairpin is the contact partner of the N-terminal tail to open/close the gate of the binding pocket. Figure S12a demonstrates that the N-terminal tail fluctuated widely with varying its orientation, which is a natural property of the intrinsically disordered segment. In contrast, the β -hairpin fluctuation is highly limited in a narrow volume as shown in figure S12b. Therefore, the gate opening/closing is induced by the large motions of the N-terminal tail (not by the fluctuation of β -hairpin).

Next, to investigate the bosentan-tail contact, we introduced a "cube-based bosentan-tail contact ratio" denoted as $\bar{c}_{b-N}(\mathbf{r}_{cube})$. Remember that the bosentan's centroid $\mathbf{r}_{CMb}$ in a snapshot was assigned to a cube at $\mathbf{r}_{cube}$, which involved $\mathbf{r}_{CMb}$. Then $\bar{c}_{b-N}(\mathbf{r}_{cube})$ is the ratio that bosentan in $\mathbf{r}_{cube}$ contacts to the N-terminal tail. The maximum and minimum of $\bar{c}_{b-N}(\mathbf{r}_{cube})$ are 1.0 and 0.0, respectively. Detailed procedure to calculate $\bar{c}_{b-N}(\mathbf{r}_{cube})$ is given in Subsection 3.3 of SI.

Figure 4a illustrates the spatial patterns of $\bar{c}_{b-N}(\mathbf{r}_{cube})$ at three contour levels: $\bar{c}_{b-N}(\mathbf{r}_{cube}) \geq 0.1$, 0.7, and 0.9. The low-contact contours, $\bar{c}_{b-N} \geq 0.1$, extended to regions distant from the gate of the binding pocket, which means that the N-terminal tail contacted bosentan even when bosentan was floating in solution. With bosentan approaching the pocket, the contact ratio increased quickly and reached $\bar{c}_{b-N} \geq 0.7$. Because the region of $\bar{c}_{b-N} \geq 0.7$ involved the density-hat region ($\rho_{CMb} \geq 0.01$; magenta-colored contour regions of figure 2), the contact may act as an attractive interaction. The high-contact region of $\bar{c}_{b-N} \geq 0.9$ (magenta-colored contours in figure 4a) was located just above the pocket not in the hETB's binding pocket. Interestingly,

$\bar{c}_{b-N}(\mathbf{r}_{cube})$ was considerably small in the pocket. The loss of the bosentan–tail contact in the pocket indicates that the role of the N-terminal tail is the recruit of bosentan into the pocket.

Remember that the gate-closing is brought by the contact of the N-terminal tail and the β -hairpin (figure S7b). In other words, the gate-opening is induced by the N-terminal tail fluctuating in solution away from the gate. Figure 4 indicate that the regions of $\bar{c}_{b-N} \geq 0.7$ are widely extended far from the binding pocket. The gate-opening/closing, the large fluctuations of the N-terminal tail and the bosentan–tail contact formation are related mutually.

Fig. 4. (a) Spatial patterns of cube-based bosentan–tail contact ratio $\bar{c}_{b-N}(\mathbf{r}_{cube})$ at three contour levels shown in inset. The “hat” region, “pocket”, and “bosentan in native complex” are also shown. (b) Spatial patterns viewed from above of the pocket. Shown structure is the native complex.

Bosentan slides along the N-terminal tail upon binding

As shown in figure 4, the cube-based bosentan–tail contact ratio depended on the relative position of bosentan to hETB in the 3D real space. Here, we quantify the bosentan’s position relative to hETB by a quantity r_{bb} , which is the distance between the bosentan’s centroid, $\mathbf{r}_{CMB,i}$, in snapshot i and that in the native complex $\mathbf{r}_{CMB}^{ref}$: $r_{bb} = |\mathbf{r}_{CMB,i} - \mathbf{r}_{CMB}^{ref}|$. Figure S7c illustrates r_{bb} in a snapshot. The smaller the r_{bb} , the closer the bosentan to its natively binding position. Then, we divided the r_{bb} axis into thirteen slices with thickness of 5 Å, which are denoted as Δ_k ($k = 1, \dots, 13$): The lower and upper boundaries of Δ_k are given as $5(k-1)$ and $5k$, respectively (unit is Å). Table S2 presents the slice range exactly. As discussed in Section 7 of SI and figure S13, the boundary between slices Δ_4 and Δ_5 coincides well to the gate of the binding pocket,

which was introduced as the boundary between the blue-colored contour region ($\rho_{CMb}(\mathbf{r}_{cube}) \geq 0.1$) and the magenta-colored contour region ($\rho_{CMb}(\mathbf{r}_{cube}) \geq 0.01$) in figure 2. By this agreement, four slices $\Delta_1, \dots, \Delta_4$ are in the binding pocket, and slices $\Delta_5, \dots, \Delta_{13}$ are outside the pocket.

Figures S14–S17 present conformations picked from some slices and Section 8 of SI explains the features of those picked conformations. These figures visually propose that the contact site on the N-terminal tail to bosentan tends to move with r_{bb} moving. To check this tendency quantitatively, we introduced a “residue-based bosentan–tail contact ratio” $\rho_{cnt}^{\Delta_k}(j)$, which is the contact ratio of residue j of the N-terminal tail to bosentan in a slice Δ_k . Then, we converted $\rho_{cnt}^{\Delta_k}(j)$ to a potential of mean force (PMF) $F^{\Delta_k}(j)$ (equation S14 of Section 9 of SI). Imagine a 2D plane constructed by the residue ordinal number j and the slice Δ_k . The lower the $F^{\Delta_k}(j)$ at a site $[j, \Delta_k]$ in the 2D plane, the stabler the contact.

Figure 5 demonstrates $F^{\Delta_k}(j)$ in the 2D plane as the free-energy landscape. This figure indicates existence of two free-energy basins, which are denoted by H_1 and H_2 in the figure. We discuss the binding process assuming that bosentan is approaching hETB from a far solvent region to the binding pocket. In Δ_{13} ($60 \text{ \AA} \leq r_{bb} < 65 \text{ \AA}$), the N-terminal tail does not touch bosentan: $\rho_{cnt}^{\Delta_{13}}(j) = 0$ ($\forall j$). In Δ_{12} ($55 \text{ \AA} \leq r_{bb} < 60 \text{ \AA}$), bosentan first touch Ace79 of the N-terminal tail. Because the system moves generally toward a low free-energy region, possible transition paths are shown by black arrows in figure 5, which lead the conformation to basin H_1 . Those motions make r_{bb} decreased and the bosentan–tail contact firm, resulting in decrement of the free energy.

With decreasing Δ_k further, the stable region moved from the basin H_1 to H_2 , that is, the system moves from the tip region to the root region of the N-terminal tail. This move is natural because H_2 is lower than H_1 . Remember that the boundary between Δ_5 and Δ_4 corresponds to the gate of binding pocket. Conformational changes from H_1 to H_2 can occur via multiple pathways indicated by the red-colored arrows in figure 5. Because the sites, to which the red arrows are assigned, had a larger free energy than H_1 and H_2 do, the H_1 - H_2 basin switch requires overcoming the free-energy barrier.

In Δ_2 and Δ_1 , which are located at the bottom of the binding pocket, the bosentan–tail contact is disrupted. This does not mean that the complex itself is dissociated but that the complex is stabilized without the bosentan–tail contact, because the native-like complex conformations are stabler than the other state as shown in figures 2 and 3. We also observed a specific contact at Arg83 in Δ_2 and Δ_1 . However, this specific contact is a weak interaction because the depth of free energy at Arg83 is shallow (figure 5).

Fig. 5. Residue-based bosentan–tail contact ratio, $\rho_{cnt}^{\Delta_k}(j)$, converted to $PMF: F^{\Delta_k}(j) = -RT \ln[\rho_{cnt}^{\Delta_k}(j)]$ (see Section 9 of SI), where j is residue ordinal number (x-axis) of hETB’s N-terminal tail, and Δ_k is the k -th r_{bb} slice (y-axis), where the lower and upper boundaries are $5(k-1)$ and $5k$, respectively. Unit is Å. PMF value is given by color (see color bar). White is assigned to unsampled regions or regions whose PMF are larger than 5 kcal / mol. Sequence of the N-terminal tail is: Ace79, Ser80, Pro81, Pro82, Arg83, Thr84, Ile85, Ser86, Pro87, Pro88, Pro89; Ace is the acetyl group introduced to cap the N-terminal tail tip. Δ_k is range of r_{bb} slice given in Table S2. Region H_1 ranges from Ser80 to Arg83 in x-axis and from Δ_7 to Δ_5 in y-axis. H_2 does from Ile85 to Pro89 and from Δ_4 to Δ_3 . Arrows are mentioned in text.

Bosentan–tail non-specific contacts

Figure 5 does not show whether the bosentan–tail contact is specific or non-specific. To investigate the bosentan’s atom specificity, we calculated an “atom–residue contact ratio” $\theta_{cnt}^{\Delta_k}(i, j)$ which is the contact ratio of a heavy atom i of bosentan contacting to residue j of the N-terminal tail in each slice Δ_k . The exact definition of $\theta_{cnt}^{\Delta_k}(i, j)$ is presented in Section 9 of SI. Figure 6 illustrates $\theta_{cnt}^{\Delta_k}(i, j)$ in each slice. The normalization is done for each panel of figure 6 to highlight the atom-specificity (color graduation) in each slice. Thus, the existence of basins H_1 or H_2 is not recognized in figure 6.

Figure 6a demonstrates that a few atoms of bosentan contacts rarely to Ace79 of the N-terminal tail in Δ_{12} . Figures 6b–d indicate that all the heavy atoms of bosentan contact non-specifically to the tip of the N-terminal tail in Δ_{11} – Δ_9 . Figures 6e–h show

again that bosentan's contact is non-specific in $\Delta_8 - \Delta_5$, where all residues of the N-terminal tail contact to bosentan. Up to here ($\Delta_{12} - \Delta_5$), bosentan is outside of the binding pocket. The non-specific contact is still found in Δ_4 (figure 6i) and Δ_3 (figure 6j), although Δ_4 and Δ_3 are in the binding pocket.

In Δ_2 and Δ_1 (figures 6k and l, respectively), contrarily, Arg83 of the N-terminal tail contacts specifically to atoms 1–13 of bosentan (see figure 6m). This specificity is consistent with figure 5. Note that the atoms 1–10 and 11–13 are hydrophobic and hydrophilic, respectively. Figure S18 demonstrates some conformations that involves the specific bosentan–tail contacts. Generally, the hydrophobic sidechain stem of Arg83 of the tail and the hydrophobic atoms 1–10 of bosentan contact, and the nitrogen atoms of the sidechain tip of Arg83 and oxygen atoms of sulfonamide of bosentan do.

As shown in figures 4 and 5, the bosentan–tail contact tends to be dissociated when bosentan proceeds to the bottom of the binding pocket (Δ_1 and Δ_2). Thus, in many snapshots in Δ_1 and Δ_2 , the N-terminal tail does not contact to bosentan. If the bosentan–tail contact has a role to capture bosentan in solution, we consider that the contact in the binding pocket has no role for molecular binding.

Fig. 6. Atom–residue contact ratio $\theta_{cnt}^{\Delta_k}(i, j)$, which is contact ratio of heavy atom i of bosentan to residue j of N-terminal tail in each slice Δ_k . Its exact definition is given in equation S15 of SI. Panels **a**, ..., **l** plot $\theta_{cnt}^{\Delta_{12}}(i, j), \dots, \theta_{cnt}^{\Delta_1}(i, j)$, respectively.

Fly-casting with ligand sliding

Here we check whether the bosentan–tail contact acts as an attractive interaction. First, we calculated a probability distribution function $P_{\Delta_k}(r_{b-N})$ regarding the minimum heavy atomic distance, r_{b-N} , from bosentan to the N-terminal tail for snapshots, which are involved in a slice Δ_k . Then, PMF regarding r_{b-N} is defined as: $PMF_{\Delta_k}(r_{b-N}) = -RT \ln [P_{\Delta_k}(r_{b-N})]$.

Figure 7 presents PMF for each slice. Apparently, the bottom of the lowest free-energy basin is assigned to a noncontact state in slices Δ_{12} – Δ_{10} ($r_{bb} \geq 45 \text{ \AA}$) (figures 7a–c). PMF for Δ_{13} is not shown because a bosentan–tail contact was not detected in Δ_{13} . In contrast, the lowest free-energy basin is assigned to a contact state ($r_{b-N} \approx 3.5 \text{ \AA}$) in Δ_9 – Δ_2 (figures 7d–k), which means that the bosentan–tail contact acts effectively as an attractive interaction in the range of $45 \text{ \AA} > r_{bb} \geq 5 \text{ \AA}$.

Above analysis showed that a bosentan–tail attractive interaction exerted from Δ_9 ($45 \text{ \AA} = r_{bb}$) between the N-terminal tail and bosentan. Together with the residue sliding from the tip side to the root side with decreasing r_{bb} , we call this binding mechanism “fly-casting with ligand sliding”.

Fig. 7. Potential of mean force $PMF_{\Delta_k}(r_{b-N})$ regarding bosentan–tail distance r_{b-N} in slice Δ_k ($k = 1, \dots, 12$). $PMF_{\Delta_k}(r_{b-N})$ is defined in main text. Panels **a**, ..., **l** plot $PMF_{\Delta_{12}}(r_{b-N}), \dots, PMF_{\Delta_1}(r_{b-N})$, respectively. $PMF_{\Delta_{13}}(r_{b-N})$ is omitted because $PMF_{\Delta_{13}}(r_{b-N})$ is similar with $PMF_{\Delta_{12}}(r_{b-N})$.

Bosentan–membrane Contact

Bosentan occasionally contacts to membrane (figures S14e, S15d and S15e). To investigate the bosentan–membrane contact, we introduced a quantity, a “bosentan–membrane contact ratio” $\bar{c}_{b-m}(\mathbf{r}_{cube})$, which is calculated with a similar procedure with $\bar{c}_{b-N}(\mathbf{r}_{cube})$: See Subsection 3.3 of SI for details of the calculation procedure. The

bosentan–membrane contact was formed commonly on the surface of the membrane (figure S19). Some experiments^{44,45} suggest a possibility that the ligand sinks and diffuses in the lipid bilayer to reach the binding pocket of GPCR rather than approaching the pocket from the solvent side. In the present simulation, however, we could not find such a ligand diffusive motion in membrane, whereas we observed bosentan sank in membrane (Figure S14e). Thus, we conclude that bosentan mainly approach the pocket from solvent. Because membrane prevents bosentan from moving to cytoplasmic side, the membrane has a role to restrict the phase volume for bosentan to move. This volume restriction may facilitate binding of bosentan to the binding pocket.⁴⁶

Orientational selection when bosentan entering the binding pocket

We investigated the molecular orientation of bosentan around hETB. First, we defined the bosentan’s molecular orientation by two vectors (orientation vectors) $\mathbf{v}_\downarrow$ and $\mathbf{v}_\leftarrow$, which are fixed in the core region of bosentan (figure S20a), which are approximately perpendicular to each other (see figure 3a for exact definition of the core region). Then, we normalized the orientation vectors as $|\mathbf{v}_\alpha| = 1$ ($\alpha = \leftarrow$ or $\downarrow$). Because the bosentan’s core region is structurally stiff, the two orientation vectors are convenient quantities to define the bosentan’s molecular orientation. We note that in general two nonparallel vectors can specify the orientation of a plane (the central ring of the core region) in the 3D real space. Figure S20b exemplifies the vectors in the native complex and in a snapshot. Next, we assigned the orientation vectors to a cube $\mathbf{r}_{cube}$, in which the bosentan’s centroid was detected. Then, Ω in equation S7 of Subsection 3.4 of SI is replaced by $\mathbf{v}_\alpha$, and the resultant quantity $\bar{\Omega}(\mathbf{r}_{cube})$ is referred to as $\bar{\mathbf{v}}_\alpha(\mathbf{r}_{cube})$ in the equation. We named $\bar{\mathbf{v}}_\alpha(\mathbf{r}_{cube})$ as an “averaged orientation vector”. Because $\mathbf{v}_\alpha$ in a snapshot was normalized, the maximum of $|\bar{\mathbf{v}}_\alpha|$ is unity when all the vectors detected in the cube $\mathbf{r}_{cube}$ are perfectly ordered, and the minimum is 0.0 when the vectors are completely randomized.

Figures 8a and b illustrate the spatial patterns of $\bar{\mathbf{v}}_\alpha(\mathbf{r}_{cube})$, in which orientation vectors with relatively large norms ($|\bar{\mathbf{v}}_\alpha| > 0.3$) are shown. Importantly, the large-norm vectors were found mainly in the binding pocket. Therefore, the bosentan’s orientation is relatively ordered in the pocket whereas it is randomized outside the pocket. This result suggests that a configurational entropy of bosentan decreases quickly when bosentan passes the gate of the binding pocket. Remember that the free-energy basin of $F^{\Delta k}(j)$ also switched quickly from H_1 to H_2 at the gate (figure 5). Thermodynamically the decrease of entropy results in increase of free energy. This entropic decrease is compensated by another thermodynamic factor, otherwise, the density $\rho_{CMB}(\mathbf{r}_{cube})$

should decrease in the binding pocket. We show in the next section that intermolecular native contacts (attractive electrostatic interactions) are formed in the binding pocket. This can enthalpically compensate the free-energetic disadvantage by the entropy decrement brought by the loss of entropy.

Figures 8a and b, however, do not necessarily mean that the bosentan's orientations are like those in the native-complex position. Then, we check the orientation similarity as follows: First, we calculated $\mathbf{v}_{\leftarrow}$ and $\mathbf{v}_{\downarrow}$ for the native complex, which are referred to as $\mathbf{v}_{\leftarrow}^{(ref)}$ or $\mathbf{v}_{\downarrow}^{(ref)}$, respectively. Next, we calculated a scalar product as:

$SP_{\alpha}(\mathbf{r}_{cube}) = \mathbf{e}_{\bar{\mathbf{v}}_{\alpha}}(\mathbf{r}_{cube}) \cdot \mathbf{v}_{\alpha}^{(ref)}$, where $\mathbf{e}_{\bar{\mathbf{v}}_{\alpha}}(\mathbf{r}_{cube})$ is a unit vector parallel to

$\bar{\mathbf{v}}_{\alpha}(\mathbf{r}_{cube})$. If $\mathbf{e}_{\bar{\mathbf{v}}_{\alpha}}(\mathbf{r}_{cube})$ is exactly parallel to $\mathbf{v}_{\alpha}^{(ref)}$, then $SP_{\alpha}(\mathbf{r}_{cube}) = 1$, and if

$\bar{\mathbf{v}}_{\alpha}(\mathbf{r}_{cube})$ is exactly anti-parallel to $\mathbf{v}_{\alpha}^{(ref)}$, then $SP_{\alpha}(\mathbf{r}_{cube}) = -1$. If $\mathbf{e}_{\bar{\mathbf{v}}_{\alpha}}(\mathbf{r}_{cube})$ and

$\mathbf{v}_{\alpha}^{(ref)}$ are perpendicular to each other, then $SP_{\alpha}(\mathbf{r}_{cube}) = 0$. The spatial pattern of $SP_{\alpha}(\mathbf{r}_{cube})$ is a measure to assess the similarity between $\bar{\mathbf{v}}_{\alpha}(\mathbf{r}_{cube})$ and $\mathbf{v}_{\alpha}^{(ref)}$.

Fig. 8. Spatial patterns of averaged orientation vectors (a) $\bar{\mathbf{v}}_{\leftarrow}(\mathbf{r}_{cube})$ and (b) $\bar{\mathbf{v}}_{\downarrow}(\mathbf{r}_{cube})$. Vector is colored differently by its norm $|\bar{\mathbf{v}}_{\alpha}(\mathbf{r}_{cube})|$ ($\alpha = \leftarrow$ or $\downarrow$) (see inset). Vectors with $|\bar{\mathbf{v}}_{\alpha}(\mathbf{r}_{cube})| < 0.3$ are

omitted. (c) Spatial patterns of scalar product $SP_{\leftarrow}(\mathbf{r}_{cube}) (= \mathbf{e}_{\bar{\mathbf{v}}_{\leftarrow}}(\mathbf{r}_{cube}) \cdot \mathbf{v}_{\leftarrow}^{(ref)})$ and (d) those of $SP_{\downarrow}(\mathbf{r}_{cube}) (= \mathbf{e}_{\bar{\mathbf{v}}_{\downarrow}}(\mathbf{r}_{cube}) \cdot \mathbf{v}_{\downarrow}^{(ref)})$, where $\mathbf{e}_{\bar{\mathbf{v}}_{\leftarrow}}(\mathbf{r}_{cube})$ and $\mathbf{v}_{\leftarrow}^{(ref)}$ are defined in main text. Regions of $SP_{\leftarrow}(\mathbf{r}_{cube}) \geq 0.5$ and $SP_{\leftarrow}(\mathbf{r}_{cube}) \leq -0.5$ are colored differently by their values (see inset). For all panels, structure of hETB is from native complex, and small black spheres are bosentan's centroid in the native complex. Binding pocket is shown by broken-line circle in all panels.

Figures 8c demonstrates that the major volume of the binding pocket was occupied by $SP_{\leftarrow}(\mathbf{r}_{cube}) \geq 0.5$, which means that the orientations of $\bar{\mathbf{v}}_{\leftarrow}$ and $\mathbf{v}_{\leftarrow}^{(ref)}$ were similar. For $\bar{\mathbf{v}}_{\downarrow}$, a large volume of the pocket was also characterized by $SP_{\downarrow}(\mathbf{r}_{cube}) \geq 0.5$ whereas a minor part (cyan-colored contours) of the pocket was characterized by $SP_{\downarrow}(\mathbf{r}_{cube}) \leq -0.5$ (figure 8d). This minor part does not cover the bosentan's native position at all. Therefore, we conclude that bosentan enters the pocket with limited orientations advantageous for reaching the native-complex position because the pocket is too narrow for bulky bosentan to flip. We call this possible mechanism "orientational-selection."

Bosentan–hETB contacts and ring-C packing at the bottom of binding pocket

Up to here, we analyzed the approaching process of bosentan to the binding pocket of hETB. In this section, we investigate the interactions of bosentan to amino-acid residues of hETB in the binding pocket.

First, we observe the intermolecular interactions in the three X-ray structures, the bosentan–hETB¹⁰, ET1–hETB⁶, and the K8794–hETB¹⁰ complexes. We note that the same binding scheme is seen in the three complex structures (figure S21). By viewing the three complex closely, we found four inter-molecular native atomic contacts, which exist commonly in the three complex structures as attractive interactions (See Section 10 of SI and figure S22): Contact (i) between oxygen atoms in sulphonyl group (SOO) of the bosentan's sulfonamide and the nitrogen (N_{ζ}) of Lys 182, (ii) between the oxygens in bosentan's SOO and the nitrogen of Lys 273, (iii) between the oxygens in bosentan's SOO and nitrogen atoms of Arg 343, and (iv) between the nitrogen atom (N_{14}) of the bosentan's sulfonamide and the nitrogen (N_{ζ}) of Lys 182.

To analyze the inter-molecular distances of snapshots, we introduce a notation $r_{\alpha;\beta}$, which is a distance between atoms α and β belonging to the ligand (bosentan, K8794, or ET1) and hETB, respectively. Then, the four distances regarding the four native contacts are denoted as $r_{N_{14};Lys182N_{\zeta}}$, $r_{SOO;Lys182N_{\zeta}}$, $r_{SOO;Lys273N_{\zeta}}$, and $r_{SOO;Arg343N_{\eta}}$. When multiple distances were possible in $r_{\alpha;\beta}$, the minimum of the

distances was selected for $r_{\alpha;\beta}$. For instance, four distances are possible for $r_{\text{SOO};\text{Arg343N}_{\eta}}$ because two oxygen atoms and two nitrogen atoms are involved in the sulphonyl group and Arg 343, respectively. Table S3 lists the distances $r_{\alpha;\beta}$ in the X-ray structures, in which three distances are around 3 Å except for $r_{\text{SOO};\text{Lys182N}_{\zeta}}$ (= 4.56 Å) from the bosentan-hETB complex. We presume that $r_{\text{SOO};\text{Lys182N}_{\zeta}}$, $r_{\text{SOO};\text{Lys273N}_{\zeta}}$, and $r_{\text{SOO};\text{Arg343N}_{\eta}}$ are related to salt-bridges, and $r_{\text{N14};\text{Lys182N}_{\zeta}}$ may be done to either a salt-bridge or a hydrogen bond. This point is discussed later.

We judged that the heavy atoms α and β are contacting if $r_{\alpha;\beta} < 5.0$ Å is satisfied in a snapshot. The value of 5.0 Å is about 0.5 Å larger than the largest (4.56 Å) of the experimental values (Table S3). Considering thermal fluctuations of the complex structure and an error in experimental measurement (3.6 Å resolution for the bosentan-hETB complex), we added the tolerance of 0.5 Å to the inequality. Besides, a heavy atom has a radius of 2.0 Å approximately, and the remaining distance between α and β is 1.0 Å (= 5.0 Å - 2 × 2.0 Å), which does not allow a water molecule to penetrate because the diameter of a water molecule is about 3.0 Å. Therefore, the threshold of 5.0 Å is rational to check the atomic contacts.

Based on the criteria above, we calculated the number of contacts N_{cnt} ($0 \leq N_{\text{cnt}} \leq 4$) regarding the four distances for all snapshots. Then, picking snapshots falling in an r_{bb} -range of $R(r_{bb}) = [r_{bb} - \Delta r_{bb}; r_{bb} + \Delta r_{bb}]$, we calculated average of N_{cnt} for the picked snapshots, and denoted it as $\langle N_{\text{cnt}}(r_{bb}) \rangle$. Section 11 of SI presents the exact method to calculate $\langle N_{\text{cnt}}(r_{bb}) \rangle$ and its standard deviation $\langle SD_{\text{cnt}}(r_{bb}) \rangle$. Ideally, Δr_{bb} should be sufficiently small ($\Delta r_{bb} \rightarrow 0$) to calculate $\langle N_{\text{cnt}}(r_{bb}) \rangle$ exactly at r_{bb} . However, the number of snapshots in $R(r_{bb})$ decreases with decreasing Δr_{bb} because the simulation length is finite. To maintain the statistical significance in $\langle N_{\text{cnt}}(r_{bb}) \rangle$, we set $\Delta r_{bb} = 0.25$ Å for the present simulation. $\langle N_{\text{cnt}}(r_{bb}) \rangle$ is 4 if all snapshots in $R(r_{bb})$ always have four contacts and zero if all snapshots always have no contact.

Figure 9 depicts the r_{bb} dependence of $\langle N_{\text{cnt}}(r_{bb}) \rangle$, where r_{bb} starts from 0.25 Å and $\langle N_{\text{cnt}}(0.25 \text{ Å}) \rangle$ is calculated for range of $R(0.25 \text{ Å}) = [0.0 \text{ Å}; 0.5 \text{ Å}]$. $R(r_{bb} < 0.25 \text{ Å})$ is impossible to be calculated. We set the maximum of r_{bb} at 15.25 Å, where the used range is $R(15.25 \text{ Å}) = [15.0 \text{ Å}; 15.5 \text{ Å}]$. Note that $R(15.25 \text{ Å})$ is involved in the slice Δ_4 , just below the gate of the hETB's binding pocket (figure S13). Therefore, figure 9 shows that almost no atom contacts are formed at the gate: $\langle N_{\text{cnt}}(15.25 \text{ Å}) \rangle = 0.005$. $\langle N_{\text{cnt}}(r_{bb}) \rangle$ increased with decreasing r_{bb} (with bosentan approaching the native-complex position): The number of contacts exceeded the

value of 1 at $r_{bb} = 3.25 \text{ \AA}$: $\langle N_{cnt}(3.25 \text{ \AA}) \rangle = 1.03 \pm 0.85$. Subsequently, the growth rate of $\langle N_{cnt}(r_{bb}) \rangle$ rose, and $\langle N_{cnt}(r_{bb}) \rangle$ reached the maximum at $r_{bb} = 0.25 \text{ \AA}$: $\langle N_{cnt}(0.25 \text{ \AA}) \rangle = 2.74 \pm 0.79$. We further investigated formation of each contact (see Section 12 of SI). Figure S23 indicates that each native contact is formed with decreasing r_{bb} .

Because the four intermolecular native contacts act as attractive interactions (Section 10 of SI), figure 9 indicates that the complex structure is stabilized enthalpically by the native contacts when bosentan goes to the bottom of the binding pocket. We presume that the instability induced by the entropy decrement when bosentan passing the gate of the binding pocket is compensated by the formation of the native contacts.

We checked the reproducibility of the packing of ring C to the bottom of the binding pocket (see figure S21c for ring C of bosentan). For this purpose, we calculated the density map, $\rho_{CMC}(\mathbf{r}_{cube})$, of the centroid of the ring C using equation S3 of Subsection 3.1 of SI with replacing the centroid of portion X by the centroid of ring C. Figure S24 demonstrates the high-density spot ($\rho_{CMC}(\mathbf{r}_{cube}) \geq 0.8$) for ring C. Importantly, the high-density spot is localized in a region that overlaps with the ring-C position in the native complex. On the other hand, the contours of $\rho_{CMC}(\mathbf{r}_{cube}) \geq 0.4$ shows a deviation from the native-complex position. We discuss this discrepancy between the computation and the native complex later.

Fig. 9. Dependency of $\langle N_{cnt}(r_{bb}) \rangle$ on r_{bb} and its standard deviation (error bars). Lower limit of some error bars is smaller than 0 on the y-axis because the standard deviation was calculated using usual form: $[\langle N_{cnt}(r_{bb})^2 \rangle - \langle N_{cnt}(r_{bb}) \rangle^2]^2$.

Bosentan's conformations in the highest-density spot

As shown in figure 2, the bosentan's centroid in the native complex structure was located on the fringe of the computed highest-density spot (red-colored contours in Figure 2). In this section, we view the bosentan's conformations in the highest-density spot.

Remember that the highest-peak position of $P_{RDF}(RMSD_{whole}^{heavy})$ or $P_{DDF}(RMSD_{whole}^{heavy})$ of figure 3 was larger than that of $P_{RDF}(RMSD_{core}^{heavy})$ or $P_{DDF}(RMSD_{core}^{heavy})$. This suggests that the sidechains of bosentan were somewhat disordered. To examine this disorder in more detail, we collected snapshots from $3.5 \text{ \AA} \leq RMSD_{whole}^{heavy} \leq 4.0 \text{ \AA}$, which involves the highest peak of $P_{RDF}(RMSD_{whole}^{heavy})$ (Figure 3b).

Fig. 10. Snapshots sampled in the highest density spot ($\rho_{CMB}(\mathbf{r}_{cube}) = 0.5$ in figure 3). In both panels, bosentan's conformations from simulation are shown in magenta, and one in the native complex is shown in green. **(a)** Four conformations picked randomly from range of $3.5 \text{ \AA} \leq RMSD_{whole}^{heavy} < 4.0 \text{ \AA}$. Bosentan's core region and ring A are shown in figure. **(b)** A conformation with $RMSD_{whole}^{heavy} = 1.11 \text{ \AA}$. hETB in the native complex is also shown. Note that superimposition was not applied to bosentan but done to trans-membrane helices of hETB (see figure S7a).

Figure 10a displays four conformations randomly pocked from the snapshots collected above. This figure demonstrates that the conformations from the highest density spot have a similarity with the native-complex structure (the X-ray structure). Thus, we referred to the complex conformations in highest density spot as “native-like complex”. On the other hand, Figure 10a indicates that the longest sidechain of bosentan, ring A (Figure S21c), is deviated from the native-complex position with a conformational

diversity, whereas the core region was relatively converged to a position close to the native-complex position. This is why the highest-peak position of $P_{RDF}(RMSD_{whole}^{heavy})$ or $P_{DDF}(RMSD_{whole}^{heavy})$ was larger than that of $P_{RDF}(RMSD_{core}^{heavy})$ or $P_{DDF}(RMSD_{core}^{heavy})$ in figure 3. We discuss the imperfectness of ring A in the next section.

Figure 10b illustrates a snapshot with a small $RMSD_{whole}^{heavy}$ (1.11 Å). Such small $RMSD_{whole}^{heavy}$ conformations were minority in the ensemble of snapshots because there was no peak at $RMSD_{whole}^{heavy} \approx 1$ Å in figure 3. This small $RMSD_{core}^{heavy}$ was resulted because the ring A of this conformation overlapped well on that of the native-complex structure.

Fig. 10. Snapshots sampled in the highest density spot ($\rho_{CMB}(\mathbf{r}_{cube}) = 0.5$ in figure 3). In both panels, bosentan's conformations from simulation are shown in magenta, and one in the native complex is shown in green. **(a)** Four conformations picked randomly from range of $3.5 \text{ Å} \leq RMSD_{whole}^{heavy} < 4.0 \text{ Å}$. Bosentan's core region and ring A are shown in figure. **(b)** A conformation with $RMSD_{whole}^{heavy} = 1.11 \text{ Å}$. hETB in the native complex is also shown. Note that superimposition was not applied to bosentan but done to trans-membrane helices of hETB (see figure S7a).

Comparison of bosentan at the highest-density spot with that in the crystal complex structure

We note several discrepancies between bosentan in the crystal complex structure and that at the computed highest-density spot. GA-guided mD-VcMD is powerful enough to explore the ligand–receptor conformational space starting from completely unbound conformations (figure S5). However, we remark that reproduction of the native-like complex was not perfect. As discussed, the bosentan's longest sidechain (ring A; figure S21c) exhibited a structural diversity in the native-like complex conformations although the core region agreed well with the native-complex position.

We presume two possibilities for the imperfectness: Accuracy of classical-mechanical force-fields and/or treatment of the bosentan's sulfonamide. Despite lots of attempts for improving the force fields, no perfect ones have not yet obtained as well known. In future, novel force-field parameters may solve the detailed conformational inconsistency with experiment.

Regarding the second possibility, we used the bosentan's sulfonamide in a normal (neutral) form in that a hydrogen atom is attached to the nitrogen atom (N14) of the sulfonamide. In the neutral form, the electronic state of N14 is the sp^2 hybridization with planner three-chemical bonds (S-NH-C angle ~ 130 degrees). In the deprotonated form, the electronic state is also the sp^2 hybridization with planner two-chemical bonds and one electron-lone pair (S = $\ddot{N}$ – C angle ~ 131 degrees). Therefore,

we cannot distinguish the electronic state of sulfonamide by viewing the X-ray structure of bosentan. We may treat the N14 atom deprotonated because the pK_a value of bosentan is 5.1 in solution⁴⁷. However, the form of sulfonamide in the binding pocket of hETB is unknown. There is a report⁴⁸ that pK_a of sulfonamide in various compounds, which are immersed in mixture of organic solvent and water, increases largely with increasing the fraction of the organic solvent. Also, as reported⁴⁹, the electronic state of aminoacids in the protein interior may be different from that in solution. Finally, we decided to use the normal form of bosentan with protonated N14. We note that the distribution function $\rho_{bs}(r_{N14;Lys182N\zeta}; r_{bb} = 0.25 \text{ \AA})$ exhibited a remarkable peak at $r_{N14;Lys182N\zeta} = 3 \text{ \AA}$ (figure S23f), reproducing the experimental observation (Table S3).

This result supports that the currently used form of sulfonamide is plausible. To analyze further the structure diversity of the ring A, huge computations are required with varying the force-fields and the charge state of sulfonamide in future.

Conclusions

Prediction of the ligand–receptor complex structure is important for structure-based drug design, and the recent fast MD simulations have enabled the ligand–receptor docking MD simulations. This strategy is reasonable to develop a new drug (especially antagonist) because the molecular system finally reaches the most stable structure regardless of the binding pathways. However, the binding process is also important to understand the complex. The current simulation study showed that an attractive interaction acts effectively between bosentan and the N-terminal tail of hETB via the fly-casting mechanism^{17,18,19} even when bosentan is located distant from the binding pocket of hETB. Some experiments have reported that the N-terminal tails of GPCRs have important physiological functions⁵⁰ whereas the N-terminal tails of GPCRs have been studied insufficiently in contrast to the binding pocket. Because the tail is conformationally disordered, a simulation like this study is helpful for understanding the function or role of the N-terminal tail.

In general, GPCRs have a long and disordered N-terminal tail (40–200 amino-acid residues long)²¹. An MD study showed that an amino-acid substitution to the N-terminal tail of a GPCR (human β_2 -adrenergic receptor) varies the dynamics of the tail, and that the accessibility of the ligand to the binding pocket also varies by the substitution⁵¹. This study as well as our present work suggest that the ligand–receptor interaction should be studied even when the ligand is outside the binding pocket of GPCR.

The bosentan–hETB complex formation obtained from the present study is summarized as follows: First the tip-side of the N-terminal tail of hETB captures bosentan when bosentan is fluctuating in solution via nonspecific intermolecular contacts. The bosentan–capturing conformation is identified as the free-energy basin H_1 in the free-energy landscape (figure 5). Then, the ligand–sliding occurs occasionally from the tip-side to the root-side of the N-terminal tail when bosentan passing the gate of the binding pocket, which corresponds to basin switching from H_1 to the deeper basin H_2 with overcoming the inter-basin free-energy barrier. We named this mechanism “fly-casting with ligand–sliding”. Importantly, the molecular orientational variety of bosentan decreases quickly when bosentan passing the gate (figure 8). This decrement of the orientational variety has a role of screening out bosentan’s conformations whose molecular orientations are improper to reach the binding site (orientational selection). We note that this orientational selection is a variant of conformational selection²⁰. Because the decrement of the bosentan’s orientational variety results in loss of a configurational entropy of bosentan, another thermodynamic quantity should compensate the free-energetic disadvantage by the decrement of the orientational variety. Otherwise, the density of the bosentan’s centroid, $\rho_{CMB}(\mathbf{r}_{cube})$, should decrease in the binding pocket. Our results show that the intermolecular native contacts (electronically attractive interactions) between bosentan and hETB are formed in the binding pocket (figure 9). It is likely that this native contact formation is the compensation factor, and eventually, the native-complex conformation is completed.

We did not use the full-length N-terminal tail for the simulation, whereas we increased the tail length than the previous study¹²: Ten residues longer than the N-terminal residues in the X-ray crystal structure for the present study, and five residues longer than those in our previous study. The present study suggests a possibility that a longer N-terminal tail adds a new property to the binding mechanism. In fact, it has been reported that long N-terminal tails exert various physiological functions of GPCRs such as ligand recognition, receptor activation, and signaling^{22,23}. Thus, a simulation with a longer N-terminal tail is the next challenge for the ligand–GPCR simulation.

The shape of the resultant free-energy landscape (potential-of-mean-force surface) was similar between the present and previous studies (figure S10 and Figure 4A of ref. 12). However, the scale of the free energy was considerably different between the two studies: The free-energy slope from the previous study was ten times steeper than that from the present study. Previously, we inferred that the steep slope from the previous study is probably caused by a drawback in the update method of $Q_{cano}^{(M)}$ ($Q_{cano}^{(M)} \rightarrow Q_{cano}^{(M+1)}$) and that the steep slope would be weakened if two pieces of $Q_{cano}^{(M)}$, which are

assigned to neighboring zones in the RC space, are connected adequately¹². GA-mD-VcMD is such a method to connect $Q_{cano}^{(M)}$ smoothly¹³. In fact, the free-energy slope decreased in the present study.

Acknowledgements

This work was supported by JSPS KAKENHI Grant No. 21K06052 (J.H.), 20H03229 (N.K.), and JP18H05534 (H.K.), and was performed in part under the Cooperative Research Program of the Institute for Protein Research, Osaka University, CR-21-05. J.H., K.K. and K.N. acknowledge support by the HPCI System Research Project (Project IDs: hp200025, hp200063, hp200090, hp210002, hp210005, and hp210008). J.H., S.I., N.H., I.F., N.K., and Y.F. are supported by the Development of innovative drug discovery technologies for middle-sized molecules project from Japan Agency for Medical Research and Development (AMED). J.H., N.K., I.F., and Y.F. are supported by Project Focused on Developing Key Technology for Discovering and Manufacturing Drugs for Next-Generation Treatment and Diagnosis (2018-2021 and 2021-) from AMED. J.H., K.K., S.S. and H.K. were partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP21am0101106. Authors thank Dr. Isao Yasumatsu from Daiichi Sankyo RD Novare Co., Ltd (Japan) for helpful discussion. GA-mD-VcMD was performed on the TSUBAME3.0 supercomputers at the Tokyo Institute of Technology.

Author contributions

All authors contributed equally.

References

1. Syrovatkina, V., Alegre, K. O., Dey, R. & Huang, X. Y. Regulation, signaling, and physiological functions of G proteins. *J. Mol. Biol.* **428**, 3850–3868 (2016).
2. Filmore, D. It's a GPCR world. *Mod. Drug Discov.* **7**, 24–27 (2004).
3. Yanagisawa, M., Kurihara, H., Kimura, S., Tomobe, Y., Kobayashi, M., Mitsui, Y., Yazaki, Y., Goto, K. & Masaki, T. A novel potent vasoconstrictor peptide produced by vascular endothelial cells. *Nature* **332**, 411–415 (1988).
4. Arai, H., Hori, S., Aramori, I., Ohkubo, H. & Nakanishi, S. Cloning and expression of a cDNA encoding an endothelin receptor. *Nature* **348**, 730–732 (1990).

5. Sakurai, T., Yanagisawa, M., Takuwat, Y., Miyazaki, H., Kimura, S., Goto, K. & Masaki, T. Cloning of a cDNA encoding a non-isopeptide-selective subtype of the endothelin receptor. *Nature* **348**, 732–735 (1990).
6. Shihoya, W., Nishizawa, T., Okuta, A., Tani, K., Dohmae, N., Fujiyoshi, Y., Nureki, O. & Doi, T. Activation mechanism of endothelin ET_B receptor by endothelin-1. *Nature* **537**, 363–368 (2016).
7. Clozel, M., Breu, V., Burri, K., Cassal, J.-M., Fischli, W., Gray, G. A., Hirth, G., Löffler, B.-M., Müller, M., Neidhart, W. & Ramuz, H. Pathophysiological role of endothelin revealed by the first orally active endothelin receptor antagonist. *Nature* **365**, 759–761 (1993).
8. Clozel, M., Breu, V., Gray, G. A., Kalina, B., Löffler, B. M., Burri, K., Cassal, J. M., Hirth, G., Müller, M. & Neidhart, W. Pharmacological characterization of bosentan, a new potent orally active nonpeptide endothelin receptor antagonist. *J. Pharmacol. Exp. Ther.* **270**, 228–235 (1994).
9. Neidhart, W., Breu, V., Bur, D., Burri, K., Clozel, M., Hirth, G., Müller, M., Wessel, H. P. & Ramuz, H. The discovery of nonpeptide endothelin receptor antagonists. Progression towards bosentan. *Chimia* **50**, 519–524 (1996).
10. Shihoya, W., Nishizawa, T., Yamashita, K., Inoue, A., Hirata, K., Kadji, F. M. N., Okuta, A., Tani, K., Aoki, J., Fujiyoshi, Y., Doi, T. & Nureki, O. X-ray structures of endothelin ET_B receptor bound to clinical antagonist bosentan and its analog. *Nat. Struct. Mol. Biol.* **24**, 758–764 (2017).
11. Hayami, T., Higo, J., Nakamura, H. & Kasahara, K. Multidimensional virtual-system coupled canonical molecular dynamics to compute free-energy landscapes of peptide multimer assembly. *J. Comput. Chem.* **40**, 2453–2463 (2019).
12. Higo, J., Kasahara, K., Wada, M., Dasgupta, B., Kamiya, N., Hayami, T., Fukuda, I., Fukunishi, Y. & Nakamura, H. Free-energy landscape of molecular interactions between endothelin 1 and human endothelin type B receptor: Fly casting mechanism. *Protein Eng. Des. Sel.* **32**, 297–308 (2019).
13. Higo, J., Kusaka, A., Kasahara, K., Kamiya, N., Hayato, I., Qilin, X., Takahashi, T., Fukuda, I., Mori, K., Hata, Y. & Fukunishi, Y. GA-guided mD-VcMD: A genetic-algorithm-guided method for multi-dimensional virtual-system coupled molecular dynamics. *Biophysics and Physicobiology* **17**, 161–176 (2020).
14. Higo, J., Kawabata, T., Kusaka, A., Kasahara, K., Kamiya, N., Fukuda, I., Mori, K., Hata, Y., Fukunishi, Y. & Nakamura, H. Molecular interaction mechanism of a 14-3-3 protein with a phosphorylated peptide elucidated by enhanced conformational sampling. *J. Chem. Inf. Model.*, **60**, 4867–4880 (2020).

15. Higo, J., Takashima, H., Fukunishi, Y. & Yoshimori, A. Generalized-ensemble method study: A helix-mimetic compound inhibits protein-protein interaction by long-range and short-range intermolecular interactions. *J. Comput. Chem.* **42**, 956–969 (2021).
16. Hayami, T., Kamiya, N., Kasahara, K., Kawabata, T., Kurita, J.-I., Fukunishi, Y., Nishimura, Y., Nakamura, H. & Higo, J. Difference of binding modes among three ligands to a receptor mSin3B corresponding to their inhibitory activities. *Sci. Rep.* **11**, 6178 (2021).
17. Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. USA* **97**, 8868–8873 (2000).
18. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
19. Arai, M. Unified understanding of folding and binding mechanisms of globular and intrinsically disordered proteins. *Biophys. Rev.* **10**, 163–181 (2018).
20. Spolar, R. S. & Record, M. T. Jr. Coupling of local folding to site-specific binding of proteins to DNA. *Science* **263**, 777–784 (1994).
21. Wallin, E. & von Heijne, G. Properties of N-terminal tails in G-protein coupled receptors: a statistical study. *Prot. Eng. Des. Sel.* **8**, 693–698 (1995).
22. Ersoy, B. A., Pardo, L., Zhang, S., Thompson, D. A., Millhauser, G., Govaerts, C. & Vaisse, C. Mechanism of N-terminal modulation of activity at the melanocortin-4 receptor GPCR. *Nat. Chem. Biol.* **8**, 725–730 (2012).
23. Uddin, M. S., Hauser, M., Naider, F. & Becker, J. M. The N-terminus of the yeast G protein-coupled receptor Ste2p plays critical roles in surface expression, signaling, and negative regulation. *Biochimica et Biophysica Acta* **1858**, 715–724 (2016).
24. Higo, J., Kamiya, N., Sugihara, T., Yonezawa, Y. & Nakamura, H. Verifying trivial parallelization of multicanonical molecular dynamics for conformational sampling of a polypeptide in explicit water. *Chem. Phys. Lett.* **473**, 326–329 (2009).
25. Ikebe, J., Umezawa, K., Kamiya, N., Sugihara, T., Yonezawa, Y., Takano, Y., Nakamura, H. & Higo, J. Theory for trivial trajectory parallelization of multicanonical molecular dynamics and application to a polypeptide in water. *J. Comput. Chem.* **32**, 1286–1297 (2011).
26. Kasahara, K., Ma, B., Goto, K., Dasgupta, B., Higo, J., Fukuda, I., Mashimo, T., Akiyama, Y. & Nakamura, H. myPresto/Omegagene: A GPU-accelerated molecular dynamics simulator tailored for enhanced conformational sampling methods with a non-ewald electrostatic scheme. *Biophys. Physicobiol.* **13**, 209–216 (2016).

27. Kutzner, C., Páll, S., Fechner, M., Esztermann, A., de Groot B. L. & Grubmüller, H. More bang for your buck: Improved use of GPU nodes for GROMACS 2018. *J. Comput. Chem.* **40**, 2418–2431 (2019).
28. Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. & Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **180**, 1961–1972 (2009).
29. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
30. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods, *J. Chem. Phys.* **81**, 511–519 (1984).
31. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions, *Phys. Rev. A* **31**, 1695–1697 (1985).
32. Kamiya, N., Fukuda, I. & Nakamura, H. Application of zero-dipole summation method to molecular dynamics simulations of a membrane protein system. *Chem. Phys. Lett.*, **568–569**, 26–32 (2013).
33. Fukuda, I., Kamiya, N., Yonezawa, Y. & Nakamura, H. Simple and accurate scheme to compute electrostatic interaction: Zero-dipole summation technique for molecular system and application to bulk water. *J. Chem. Phys.* **137**, 054314 (2012).
34. Fukuda, I., Yonezawa, Y. & Nakamura, H. Molecular dynamics scheme for precise estimation of electrostatic interaction via zero-dipole summation principle. *J. Chem. Phys.* **134**, 164107 (2011).
35. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct. Funct. Genet.* **78**, 1950–1958 (2010).
36. Dickson, C. J., Madej, B. D., Skjevik, Å. A., Betz, R. M., Teigen, K., Gould, I. R. & Walker, R. C. Lipid14: The Amber lipid force field. *J. Chem. Theory Comput.* **10**, 865–879 (2014).
37. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–936 (1983).
38. Joung, I. S. & Cheatham, III, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
39. Frisch, M. J. *et al.* Gaussian 09, Revision D.01.

40. Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).
41. Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
42. Wacker, D., Wang, S., McCorvy, J. M., Betz, R. M., Venkatakrishnan, A. J., Levit, A., Lansu, K., Schools, Z. L., Che, T., Nichols, D. E., Shoichet, B. K., Dror, R. O. & Roth, B. L. Crystal structure of an LSD-bound human serotonin receptor. *Cell* **168**, 377–389 (2017).
43. Che, T., English, J., Krumm, B. E., Kim, K., Pardon, E., Olsen, R. H. J., Wang, S., Zhang, S., Diberto, J. F., Sciaky, N., Carroll, F. I., Steyaert, J., Wacker, D. & Roth, B. L. Nanobody-enabled monitoring of kappa opioid receptor states. *Nat. Commun.* **11**, 1145–1145 (2020).
44. Szlenk, C. T., Gc, J. B. & Natesan, S. Does the lipid bilayer orchestrate access and binding of ligands to transmembrane orthosteric/allosteric sites of G protein-coupled receptors? *Mol. Pharmacol.* **96**, 527–541 (2019).
45. Toyoda, Y., Morimoto, K., Suno, R., Horita, S., Yamashita, K., Hirata, K., Sekiguchi, Y., Yasuda, S., Shiroishi, M., Shimizu, T., Urushibata, Y., Kajiwara, Y., Inazumi, T., Hotta, Y., Asada, H., Nakane, T., Shiimura, Y., Nakagita, T., Tsuge, K., Yoshida, S., Kuribara, T., Hosoya, T., Sugimoto, Y., Nomura, N., Sato, M., Hirokawa, T., Kinoshita, M., Murata, T., Takayama, K., Yamamoto, M., Narumiya, S., Iwata, S. & Kobayashi, T. Ligand binding to human prostaglandin E receptor EP₄ at the lipid-bilayer interface. *Nat. Chem. Biol.* **15**, 18–26 (2019).
46. McCloskey, M. A. & Poo, M. M. Rates of membrane-associated reactions: reduction of dimensionality revisited. *J. Cell Biol.* **102**, 88–96 (1986).
47. Gatfield, J., Grandjean, C. M., Bur, D., Bolli, M. H. & Nayler, O. Distinct ETA receptor binding mode of macitentan as determined by site directed mutagenesis. *PLoS ONE* **9**, e107809 (2014).
48. Şanlı, N., Şanlı, S., Özkan, G. & Denizlic, A. Determination of pK_a values of some sulfonamides by LC and LC-PDA methods in acetonitrile-water binary mixtures. *J. Braz. Chem. Soc.* **21**, 1952–1960 (2010).
49. Isom, D. G., Castañeda, C. A., Cannon, B. R. & García-Moreno, E. B. Large shifts in pK_a values of lysine residues buried inside a protein. *Proc. Natl. Acad. Sci. USA* **108**, 5260–5265 (2011).
50. Coleman, J. L. J., Ngo, T. & Smith, N. J. The G protein-coupled receptor N-terminus

and receptor signalling: N-tering a new era. *Cell. Signal.* **33**, 1–9 (2017).

51. Shahane, G., Parsania, C., Sengupta, D. & Joshi, M. Molecular insights into the dynamics of pharmacogenetically important N-terminal variants of the human β_2 -adrenergic receptor. *PLoS Comput. Biol.* **10**, e1004006 (2014).