ACS Chemical Biology

Unraveling the Conformational Landscape of Ligand Binding to Glucose/Galactose-Binding Protein by Paramagnetic NMR and MD Simulations

Luca Unione,[†] Gabriel Ortega,[†] Alvaro Mallagaray,[‡] Francisco Corzana,[§] Javier Pérez-Castells,[‖] Angeles Canales,[⊥] Jesús Jiménez-Barbero,[*,†,#,∇] and Oscar Millet[*,†]

[†]Molecular Recognition and Host−Pathogen Interactions, CICbioGUNE, Bizkaia Technology Park, Building 801 A, 48170 Derio, Spain

[‡]Institute of Chemistry, Center for Structural and Cell Biology in Medicine (CSCM), University of Luebeck, Ratzeburger Allee 160, 23538 Luebeck, Germany

[§]Department of Chemistry, University of La Rioja, 26006 Logroño, Spain

[‖]Facultad de Farmacia, Dpto. Química y Bioquímica, Universidad San Pablo CEU, Urb. Montepríncipe, ctra., Boadilla km 5,300 Boadilla del Monte, 28668 Madrid, Spain

[⊥]Department of Química Orgánica I, Fac. C. C. Químicas, Universidad Complutense de Madrid, Avd. Complutense s/n, 28040 Madrid, Spain

[#]Ikerbasque, Basque Foundation for Science, Maria Diaz de Haro 13, 48009 Bilbao, Spain

[∇]Departament of Organic Chemistry II, Faculty of Science & Technology, University of the Basque Country, 48940 Leioa, Bizkaia Spain

S *Supporting Information*

ABSTRACT: Protein dynamics related to function can nowadays be structurally well characterized (i.e., instances obtained by high resolution structures), but they are still ill-defined energetically, and the energy landscapes are only accessible computationally. This is the case for glucose−galactose binding protein (GGBP), where the crystal structures of the apo and holo states provide structural information for the domain rearrangement upon ligand binding, while the time scale and the energetic determinants for such concerted dynamics have been so far elusive. Here, we use GGBP as a paradigm to define a functional conformational landscape, both structurally and energetically, by using an innovative combination of paramagnetic NMR experiments and MD simulations. Anisotropic NMR parameters induced by self-alignment of paramagnetic metal ions was used to characterize the ensemble of conformations adopted by the protein in solution while the rate of interconversion between conformations was elucidated by long molecular dynamics simulation on two states of GGBP, the closed-liganded (*holo_cl*) and open-unloaded (*apo_op*) states. Our results demonstrate that, in its apo state, the protein coexists between open-like (68%) and closed-like (32%) conformations, with an exchange rate around 25 ns. Despite such conformational heterogeneity, the presence of the ligand is the ultimate driving force to unbalance the equilibrium toward the *holo_cl* form, in a mechanism largely governed by a conformational selection mechanism.
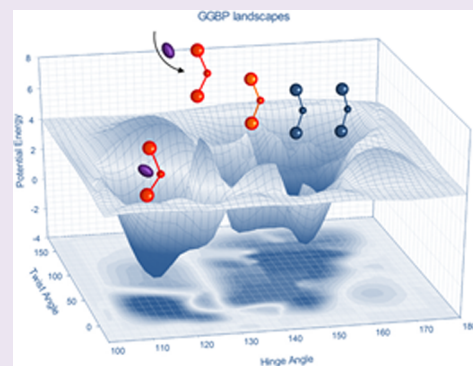
Protein function arises from the delicate interplay among structure, molecular recognition features, and dynamics, but unraveling such contributions is often elusive. The periplasmic binding protein family (PBPs) represents a paradigm for describing functional conformational changes in flexible proteins.[1] In Gram-negative bacteria, PBPs selectively recognize and actively transport various nutrients across the inner membrane. The family is composed of about 100 members, classified according to the recognized ligand: amino acids, carbohydrates, oxyanions, and vitamins.[2] Almost all of them share a common structural fold consisting of two globular Rossman fold domains connected by three short linkers, thus suggesting interdomain flexibility.[3−5] This hypothesis is further supported by the different interdomain orientations found in the X-ray structures. For instance, one of the paradigmatic glycan-binding protein family, the glucose/galactose binding proteins (GGBP) from different organisms have generated a plethora of crystal structures that trap the biomolecule at distinct conformational instances: some unloaded structures are canonically open (*apo_op*), while others are closed (*apo_cl*)

A   DOI: 10.1021/acschembio.6b00148
ACS Chem. Biol. XXXX, XXX, XXX−XXX

and resemble the holo-ligand-bound state (*holo_cl*).[6−10] In the past few years, segmental interdomain reorientations in periplasmic binding proteins have been extensively investigated by solution NMR spectroscopy.[11] Clore and co-workers have demonstrated that a conformational selection process undergoes an open-to-closed transition in MBP,[12] while Tjandra and co-workers have shown that an induced fit mechanism well describes the open−closed transition of another PBP, glutamine-binding protein.[13] In a comparative NMR study of GGBP and the structurally homologous ribose binding protein, it was shown that the (apparent) ligand affinity can be modulated by redesigning the flexible hinge region, thus emphasizing the functional role of interdomain dynamics.[14] However, the time scale and amplitude of these motions are experimentally ill-defined for all the investigated cases. Molecular dynamics simulations have also been widely used to characterize the conformational landscape of PBPs.[3] For instance, advanced sampling techniques have been used to study the allosteric equilibrium of the ribose-binding protein,[15] while accelerated MD simulations provided a detailed picture of the transition between the open and partially closed states in MBP.[16] Moreover, these proteins have been the target of intense studies in protein engineering,[17] and the computational redesign of PBPs to build up nanobiosensors have raised great expectations.[18,19] For instance, Daunert and co-workers have proposed GGBP as a possible biosensor of glucose in blood.[20] However, ironically, the main limitation of the method is the high affinity for the substrate (nM range). Then, despite the extensive use of MD simulations in the study of PBPs, integrative approaches of protein design with experimental data are still largely unedited.

Here, we propose an integrated approach by combining experimental NMR data with molecular dynamics simulations to quantitatively characterize interdomain dynamics in GGBP. First, pseudocontact shifts (PCSs) and residual dipolar couplings (RDCs) allowed for disentangling the population distribution of conformers in the open−closed transition that GGBP undergoes. Next, the time scale for the open−closed transition is defined by detailed molecular dynamics simulations. Finally, the energy barrier in the protein landscape has been estimated using nonequilibrium molecular dynamics calculations. Our results demonstrate that, in its *apo* state, the protein coexists between the open (68%) and closed (32%) conformations. The time scale for closed−open interconversion is around 25 ns. The presence of the ligand is the driving force for closing, largely through a conformational selection mechanism.

■ RESULTS AND DISCUSSION

Self-Alignment with a Paramagnetic Tag Reveals Conformational Heterogeneity in apoGGBP. GGBP consists of two globular domains, the C-terminal domain (residues 112−254 and 297−306) and the smaller N-terminal domain (residues 3−108 and 258−291), linked by a three-strand hinge (residues 109−111, 255−258, and 292−296) (Figure 1, panel A). The β anomer of glucose binds to GGBP through an extensive network of hydrogen bonds and CH-π interactions with high affinity and specificity (K_D = 290 nM at 37 °C and pH 7.0) into the cleft near the hinge region, invoking a large conformational change from the open unbound (*apo_op*) to closed bound (*holo_cl*) state.[6] This segmental interdomain reorientation is well described by a rotation of 40° in the twist (φ) angle accompanied by a 23° rotation in the
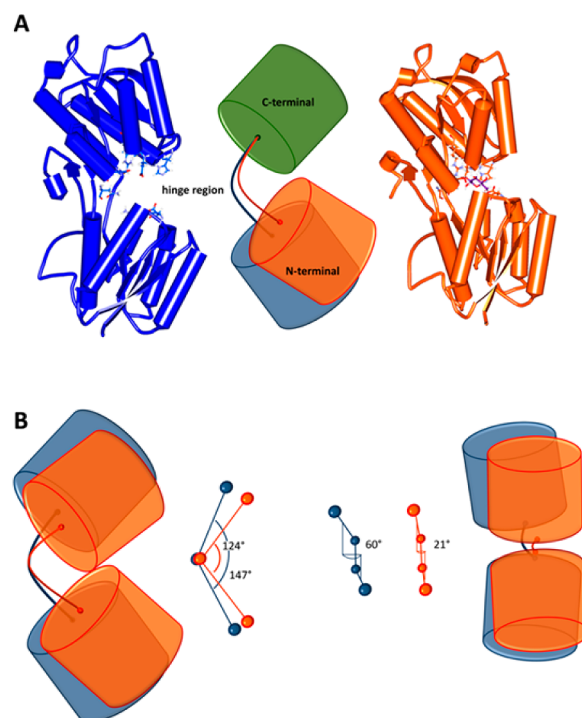

Figure 1. (A) Crystal structures of *apo_op*GGBP (left, 2FW0) and *holo_cl*GGBP (right, 2FVY). The bound β anomer of D-glucopyranose and residues Asp14, Asn91, His152, Asp154, Arg158, Asn211, Asp 236, and Asn256 forming stabilizing H-bonds with the ligand are drawn as stick models (ligand in violet). Inset: schematic representation of globular domains (cylinders) and the hinge region (lines). The difference in N-terminal domain position highlights the difference in closure angle, according to X-ray structures. (B) Domain reorientation of GGBP. Left, side view illustrating hinge domain rearrangement between *apo_op* (blue) and *holo_cl* (orange) GGBP. Right, front view illustrating twist motion. Inset: the angle between the segments connecting the center of mass of the hinge region and those of the C-terminal domain and N-terminal domain is defined as hinge angle, while the center of mass of the N-terminal domain, the base of the N-terminal domain, the C-terminal domain and the base of the C-terminal domain define three segments. The dihedral angle formed by these three segments is defined as twist angle.

hinge (θ) angle[14,15] (Figure 1, panel B). Since closed unloaded structures are available for some PBPs, an open question is whether *apo*GGBP can make excursions to the *holo_cl* conformation in the absence of ligand. Structural data for GGBP in solution can be integrated by anisotropic NMR parameters that are induced by self-alignment of paramagnetic molecules. Such self-alignment has been achieved by binding paramagnetic metal ions to a small molecule chemical metal chelating tag, covalently attached to the biomolecule.[21] The synthesis of the novel tag is described in the Supporting Information. The alignment tensors for the tagged and nontagged domains have been determined by the combined use of residual dipolar couplings (RDCs) and pseudocontact shifts (PCSs). Actually, owing to the large magnetic moment of the unpaired electrons on the paramagnetic lanthanide ion, the paramagnetic effects are detectable up to large distances (>40 Å).[22] In a system comprising two or more species in rapid exchange, the observed effect is a population-weighted average of the component conformers. As a result, PCSs and RDCs provide a unique way to describe complex mixtures of translational and rotational interdomain motions, simply

aligning the tagged domain by the paramagnetic tag and determining the induced alignment on the other moieties.[23] To that end, the paramagnetic probe has been conjugated through nucleophilic substitution to an engineered cysteine residue, M182C, located in the C-terminal domain at the periphery of the interface between the two domains. Several factors were considered for the insertion of the tag molecule: (i) the minimal biorelevant mutation Cys instead of Met, (ii) an adequate distance between the tag molecule and the target binding site, so the spin-label does not perturb the backbone structure nor the ligand-binding site, and (iii) surface-accessible amino acids that experience minimum variation in chemical environment upon sugar binding (Figure S1). The NMR signals for the residues within the shell around the paramagnetic center (C182 and A181) were broadened beyond detection due to paramagnetic relaxation enhancement (PRE). Nevertheless, the chemical nature of the tag molecule (the metal is located at a distance >16 Å away from the protein backbone) allowed collecting 135 measurable PCSs (Figure 2)

Figure 2. Pseudocontact shifts (PCSs) obtained as the difference in the chemical shift of the protons signals in diamagnetic (lanthanum) and paramagnetic (dysprosium) conditions, for the observable H^N, N nuclei. Resonances of residues within the shell around the paramagnetic center were broadened beyond detection due to paramagnetic relaxation enhancement (PRE).

and 20 RDCs. Clearly, two regions of the protein orient differently with respect to the paramagnetic metal ion. One region undergoes negative chemical shift perturbation, close in space to the negative lobe of the magnetic susceptibility tensor. The other region experiences opposite changes due to its orientation toward the positive lobe of the paramagnetic metal isosurface.

To take into account the flexibility of the linker, two main conformations with staggered dihedral angles around the disulfide bond have been generated (90° and −90°). Only the −90° conformer fits well the experimental data, confirming the impossible population of the 90° conformer due to steric clashes (see Supporting Information). PCSs of H^N, N nuclei, and RDCs of H^N-N pairs for paramagnetic lanthanide (Dy^{3+}) were determined from an $^1H,^{15}N$ HSQC. A La^{3+}-loaded complex was used as diamagnetic reference, as this metal has an ionic radius similar to that of the dysprosium ion. The protein shows excellent signal dispersion in both $^1H,^{15}N$-HSQC spectra (Dy^{3+} and La^{3+}), (Figure 3). Chemical shifts between unloaded and diamagnetic metal ion loaded molecules are virtually identical, and nearly complete assignments could be obtained based on previously published data.[14] The lack of chemical shift perturbation in the GGBP backbone amide signals after metalation of the sample indicates that all the

lanthanide ions are bound to the tag and not directly to the protein, consistent with the high affinity of the tag for lanthanides (in the 10−18 M range).[24,25] Representative structures for the *apo_op* and *apo_cl* conformations (with 5° stepwise changes in the closure and twisting angles) were extracted from a molecular dynamic simulation (vide infra) and used for the alignment tensor estimation, using RDCs and PCSs as experimental restraints. The structures providing the lowest Q-factor for the tagged domain were selected and subsequently used for the prediction of the NMR parameters in the tagged-free domain. Two structures showed the lowest quality factors (Q-factor) when fitting the experimental data for the tagged domain: one representative for the *apo_op* form (θ = 145° and ϕ = 64°) and another for the *apo_cl* conformation (θ = 137° and ϕ = 20°). When analyzed independently, PCSs and RDCs (the latter to a lower extent) provided excellent fits for each tagged domain for both structures, as reflected in the range of the Q-factor values: 0.084−0.099 (PCSs) and 0.494−0.420 (RDCs).

These results demonstrate that the Tag molecule and the C-terminal domain behave as a rigid body (Figure 4, panel A and intra Q-factor in panels C−E). Interestingly, Q-factors drastically increase when both domains are included in the analysis (Figure. 4, panel B and overall Q in panels C,D). When fitting the PCS data set, the quality factor referred to the open structure rises up to 0.231, while when referred to a closed-like structure, it reports a value of 0.281. These results demonstrate that a single structure is not able to explain the experimental data set, likely because the N-terminal domain fluctuates with respect to the C-terminal. A model contemplating an average ensemble of differently populated states was tested, and a combination of the two above-mentioned conformations (68% for the *apo_op* and 32% for the *holo_cl*) provides a very good correlation with the experimental data (Figure 4, panel E). Thus, the RDC and PCS experimental data reported here on *apo*GGBP fully agree with a model where the *apo* state undergoes a rapid equilibrium between a major and a minor species, the latter one occupying a region of the conformational landscape similar to the ligand bound form (Figure 5, panel C). Our results also demonstrate the existence of the postulated dynamic equilibrium between open and partially closed apo states and gives credit to the hypothesis that large-scale domain rearrangements are already present in many two-domain periplasmic proteins.

Unraveling the Conformational Landscape of GGBP by MD Simulations. Because of the dynamic nature of *apo*GGBP, MD simulations have been employed to investigate its conformational landscape. In a first set of calculations, the experimental RDCs were included as restraints in a conjoined rigid body-torsion angle simmulated annealing followed by an MD simulation of 200 ps. (see Methods). An intermediate conformation, partially closed, complies well with the experimental restraints and can be interpreted in terms of a combination of the open (68%) and closed (32%) states. Specifically, the values for the interdomain hinge/twist angle that satisfies the NMR restraints lies around 135 ± 5°/40 ± 10° (see Figure 5). The values are similar but statistically different from the consensus conformation obtained from fitting experimental diamagnetic RDCs in weakly aligned media (θ 127°; ϕ 32°),[14] thus suggesting the existence of interdomain dynamics.

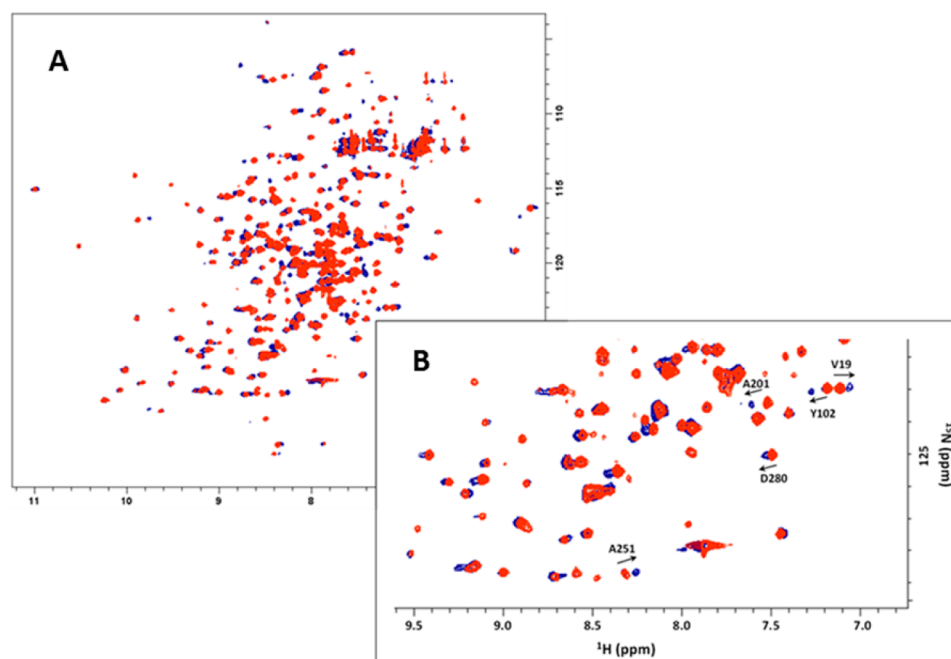To further characterize such motions, in a second calulation experimentally determined NMR constraints (i.e., PCSs and

Figure 3. (A) ^{1}H,^{15}N HSQC for GGBP loaded with diamagnetic lanthanum (orange) and paramagnetic dysprosium (blue). Spectra were recorded at 310 K and pH 7.0 in a 600 MHz NMR spectrometer. (B) Selected spectral region of the ^{1}H,^{15}N HSQC spectra.

RDCs) were not introduced to minimize biases in the conformational landscape. Two MD simulations, starting from X-ray structures of *apo_op* (2FW0) and *holo_cl* GGBP bound to glucose (2FVY), were performed with the ff10_Amber and GLYCAM_06h force fields in explicit water at 310 K to generate atomic coordinate data sets that describe the protein ensemble. Because of the expected higher dynamic complexity, the MD simulation for *apo_op* was run for 300 ns, while *holo_cl* was run for a time of 100 ns. It is important to emphasize that both free-MD simulations are long enough to explore a wide conformational space and to define the periodic dynamic behavior of GGBP (*apo* and *holo*) in solution.

After excluding the preparatory steps from the trajectory, MD runs have been analyzed in terms of the interdomain hinge and twist angles with a defined hinge position (Figure 5, panels A,B). For *apo_op* GGBP, a range of twist and hinge angles are accessible, oscillating between a partially closed conformation (θ 135°; ϕ 40°) and a widely open conformation (θ 170°; ϕ 160°), consistent with previous studies.[6] Conformational interconversion in *apo_op* GGBP involves concerted changes in both angles: the twist angle fluctuates between 30° and 160°, with equally low energy conformers, and it is always accompanied by a hinge angle oscillation between 140° and 160°. The time scale for such collective dynamics, between crest and wave of the periodic motion, is around 25 ns. As expected, MD simulation for *holo_cl* GGBP structure is characterized by minor excursions within the conformational space. The hinge and the twist angles oscillate between ±10° and ±20°, respectively, around the starting values.

The overall QRDC value estimated for the ensemble structure between the *apo_op*(X-ray) (label 4 in Figure 6) and *holo_cl*, (label 8 in Figure 6) markedly improves with respect to the single conformation ones (see Figure 4, panel C−E). The results demonstrate that the conformational behavior of GGBP in its unbound state is compatible only with a dynamic phenomenon that involves partially closed conformations suggesting that the ligand recognition event

cannot be described by pure "induced-fit" or pure "conformational selection" models.

The free energy of the conformational landscape explored by the MD simulations has been estimated from the populations by dividing the conformational space into regular intervals of (2° × 4°) hinge/twist angles. The population for each interval has been converted into the energy difference with respect to the highest populated one according to the following expression:

$$\Delta G_{\text{PA-PB}} = -RT \ln([P_A]/[P_B])$$

The free energy landsacpe for *apo_op*GGBP (Figure 5, panel C) shows a wide global minimum corresponding to an ensemble of highly variable structures in terms of twist/hinge angles. Remarkably, a variety of conformations expanding the open-to-close conformational coordinate (i.e., superopened and closed-like conformations) have a free energy excess of only 1− 2 kcal·mol^{-1} as compared to the global minimum, indicating that these structures are also accessible. As inferred from Figure 5, panels A,B, these conformations are periodically revisited, demonstrating that they are not transiently populated high energy states. However, the free energy landscape for *holo_cl*GGBP fluctuates around a global minimum, structurally corresponding to the starting conformation. Large scale conformational changes are multiple trajectory processes. Nevertheless, the two independent molecular dynamic simulations seem to energetically coalesce, defining the conformational coordinate for the most likely open-to-close trajectory.

In order to well define the energy profile associated with the open-to-close transition in GGBP, SMD have performed. This computational approach has been extensively used to calculate the free energy associated with unfolding/refolding pathways of macromolecules,[26,27] ligand−receptor binding events,[28] or DNA starching.[29] SMD employs a pulling force to cause a structural change so that different conformations may be sampled along a given pathway. To gain access to time scales
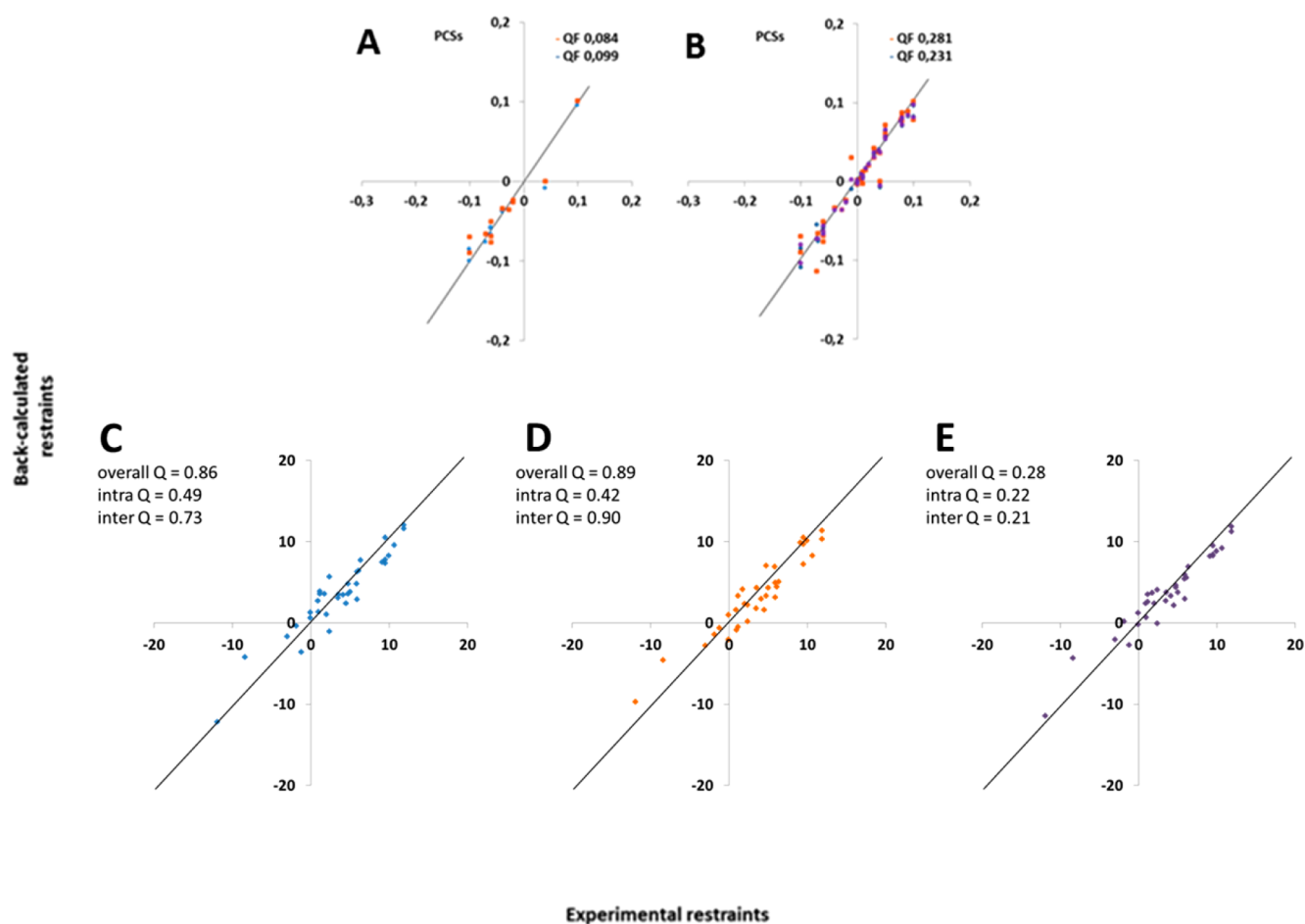
Figure 4. Correlation between experimental and back-calculated restraints relative to *apo_op*GGBP (blue), *holo_cl*GGBP (orange), and an average of 68% and 32% of both structures, respectively (violet). (A) PCSs of H^N, N nuclei belonging to only the protein tagged domain. (B) PCSs of H^N, N nuclei of the entire protein. (C) RDCs of H^N-N pairs of different selected amino acids in *apo_op*GGBP. Label 4 in Figure 5. (D) RDCs of H^N-N pairs of different selected amino acids in *holo_cl*GGBP. Label 8 in Figure 5. (E) RDCs of H^N-N pairs of different selected amino acids in *apo_cl*GGBP. This latter structure is an ensamble average structure of 68% *apo_op* and 32% *holo_cl* GGBP. The values for hinge and twist angles are 137° and 40°, respectively, and its position in the conformational space is specified in Figure 5.
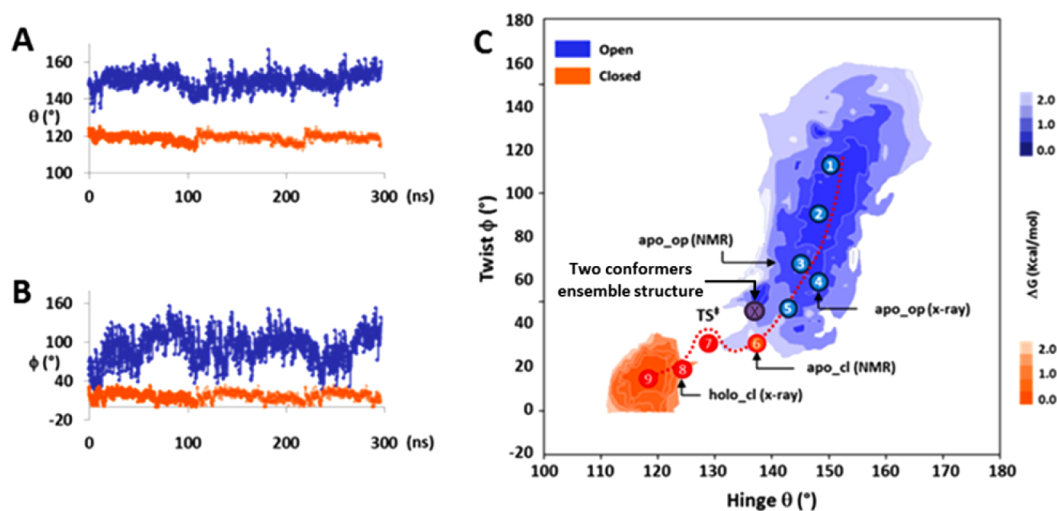

Figure 5. Analysis of interdomain hinge (A) and twist (B) angles along the MD simulations for *apo_op*GGBP (300 ns) and *holo_cl*GGBP (100 ns). The derived 200 ns of collective motion for *holo_cl*GGBP are marked with a tenuous line. (C) Free-energy landscapes of *apo_op*GGBP (blue) and *holo_cl*GGBP (orange) as a function of hinge ($\overline{\theta}$) and twist (ϕ) angles. Selected snapshot along the trajectory, red dashed line, are labeled with numbers in circles. From *apo_op* MD, 1−3 and 5,6; 4 is the X-ray structure (2FW0); 4 and 8 have been used to derive the ensemble average population from experimental NMR restraints. From *holo_cl* MD, 8,9; from steered molecular dynamic simulations (SMD), 7 is the high energy structure corresponding to the transition state.
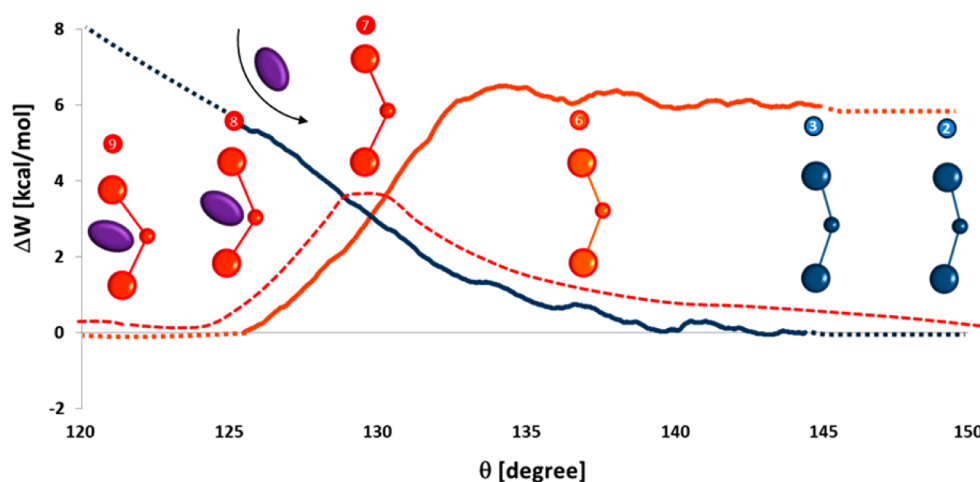
Figure 6. Potential of mean force as a function of the hinge angle θ (deg) in protein/ligand binding. The calculated energy profile relative to the *holo_cl* structure (orange) indicates that the free energy minimum for closed conformation is lower than that of the open state by roughly 6 kcal·mol^{-1}. Superimposed is the energy profile for the *apo_op* structure (blue). The result indicates that the energy of the open structure increases with the closing of the hinge angle. On top, structural change associated with interdomain closure and ligand binging. The structures correspond to snapshots along the trajectory.

that would be otherwise computationally too demanding, SMD simulations accelerate (or force) the dynamic process, thus providing the required energy associated with the induced change: while the force is executed and the motion occurs along a given coordinate, the potential energy of the system is calculated, and the potential of mean force (PFM) is related to the free energy profile of the process.[30] Even if we cannot exclude the existence of other transition coordinates for the open-to-close event, the method ensures that the work done on the system is a function of the activation energy associated with the process. Two SMD simulations have been carried out. In the first one, the value of the critical hinge angle was fixed to 124°, which corresponds to the X-ray structure determined for the *holo_cl* conformation, and a pulling force was applied to reach the *holo_op* conformation. In the second one, the starting geometry corresponds to the X-ray structure of the unbound open GGBP (*apo_op*), with a hinge angle of 145°, in order to achieve the closed unbound (*apo_cl*) structure. Figure 6, orange line plots the resistance applied by the system in opening *holo_cl*GGBP as a function of the hinge angle. According to the SMD simulations, the transition state corresponds to a hinge angle of 132°, and its height is only ≈3.5 kcal·mol^{-1} higher than the free energy of the open structure.

We can conclude that the energetic barrier for open-to-closed interconversion can be easily crossed via thermal fluctuations or with the help of water molecules that attack interdomain hydrogen bonds. Not surprisingly, the closed structure is significantly stabilized by intermolecular interactions with the ligand (hydrogen bond and CH/π interactions) that contribute with 6 kcal·mol^{-1} to the binding energy of the complex. The thick hydrogen bond network between the polar face of the sugar and the polar amino acids on both domains of GGBP and the two aromatic residues Trp181 and Phe16 on the C-terminal and N-terminal domain, respectively, stacked against the less polar faces of the sugar provide the necessary enthalpy contribution to stabilize the *holo_cl*GGBP structure. The second SMD simulation starting in the *apo_op*GGBP is also shown in Figure 6, blue line. Here, a quasi-linear energy dependence is observed in the transition

from the open to the closed states of GGBP, with an associated Gibbs free energy that is inversely proportional to the closing angle. The stability of the protein decreases at a rate of about 300 cal·mol^{-1}·deg^{-1}. Considering a range of 3 kcal·mol^{-1} accessible by thermal fluctuations at 310 K, a domain closure of 10° to 20° is expected. Then, the *apo_cl* conformation of the protein, which is almost 6 kcal·mol^{-1} higher in energy than the *apo_op*, would be inaccessible by thermal fluctuations, consistent with previous results reported for MBP.[11] The intersection between the two independent SMD define the transition state, where only the incoming ligand provides the necessary energy stabilization to shift the equilibrium in favor of the most stable ligand−protein complex.

Synergy between Structure and Dynamics Is Essential for Protein Function. Segmental dynamics in proteins is often functional and is the basis of protein allosterism. Yet, they are often very loosely characterized due to the lack of experimental tools available. Here, we demonstrate that a combination of NMR spectroscopy and MD simulations successfully unravels the energy landscape for the functional conformational coordinate of GGBP.

A single structural model unsatisfactory describes the vast conformational space covered by *apo_op*GGBP in solution. According to unrestricted MD simulations, the conformational dynamics is well characterized by large amplitude motions between hemiclosed and open structures in the nanosecond time scale. Actually, the first reported X-ray structure of *apo_*GGBP is partially closed.[6] The authors noticed the presence of the citrate ion in the binding cleft and hypothesized that they may be thermodynamically relevant for crystal growth. Our results agree, demonstrating that the sodium ion coordinates at the hinge position, thus stabilizing the hemiclosed structures in *apo_*GGBP. Importantly, these conformational interconversions are experimentally validated, while the combined use of PCSs and RDCs induced by self-alignment of paramagnetic metal provides a fine method to unveil the existence of functional closed conformations that are consistent with a 32% of the total population of *apo_*GGBP.

Albeit the limited range of closure angles for the deposited X-ray structure of *apo_op*GGBP spans between 147° and 149°,

our data suggest that excursions toward more open states are lightly permitted (150° and 170°, "superopen" structures). Such conformations have been also predicted and observed for other bidomain proteins like uroporphyrinogen III synthase,[31] and it is probably dependent on the hinge region structure and composition since it reflects the maximum torque force that this region allows. However, closed-like states are also represented, between 132° and 147°, supporting the existence of closed-like structures for *apo*GGBP as observed via X-ray crystallography.[32]

The interdomain excursions found for *apo_*GGBP are largely quenched in *holo_*GGBP, but interestingly, they are not completely abrogated (Figure 5). Actually, when bound to glucose, *holo_cl*GGBP still holds some degree of flexibility, with interdomain librations of up to 10−15° in the closure and hinge angles. According to MD simulations, such motions also fall in the nanosecond time scale, and they are consistent with the increased dynamics previously observed in the order parameter of *holo_*GGBP[14] that likely interfered on the spectral density function motional parameters.

The functional role of these dynamics naturally emerges when the energetics are introduced, and the integrative analysis of MD simulations and NMR analysis comply largely with a mainly conformational selection mechanism for which the presence of the ligand is indispensable for the open-to-closed transition. MD simulations reveal a gap in the conformational landscape between *apo_op*GGBP and *holo_cl*GGBP (125−140° hinge and 30−40° twist, Figure 5, panel C). This high energetic point corresponds to the transition state with an activation barrier ≈3.5 kcal·mol^{-1} as estimated by SMD simulations (Figure 6). Thus, ligand binding provides the energy to overcome such a barrier and shift the conformational ensemble toward *holo_cl*GGBP in a second step that agrees well with an induced fit mechanism. The open-to-closed transition for GGBP is described in a short representative video in Video 1. Remarkably, *apo_op*GGBP is predicted to unfold at low values of the closing angle due to the increase in nonpolar solvent accessible area located on the hinge region on the opposite site with respect to the ligand binding pocket. This negative term, if not balanced by the enthalpic surplus of the binding event, will result in protein unfolding. This mechanism is equivalent to the experimentally found one for maltose binding protein, where the analysis of hinge mutants with different closure angles predicted protein unfolding at low closure angles, within a similar free energy range.[11]

In conclusion, the amplitude and time scale of GGBP interdomain dynamics have been unveiled by NMR spectroscopy and detailed MD simulations. The population contribution of biofunctional relevant conformers has been determined by PCSs and RDCs induced by a paramagnetic metal ion. Subsequently, the energetic barrier in open-to-closed transition has been defined by nonequilibrium MD simulations. We conclude that the population of *apo_cl*-like conformations is essential to activate the transition to the *holo_cl* form, according to a conformational selection mechanism coupled to a final rearrangement that obeys an induced-fit kinetics process.

From a general perspective, the protocol exemplified herein can be extended to the study of a variety of molecular recognition processes in which significant molecular rearrangements take place, thus expanding the limits of the application of NMR methods to explore binding events.

■ METHODS

Protein Expression and Purification. Uniformly labeled (^{15}N) GGBP protein was overexpressed in minimum M9 media (1.5 L, purchased from CIL) containing ^{15}NH$_4$Cl as the only source for nitrogen. GGBP was purified via ion exchange chromatography using fast flow Q-Sepharose followed by ion exclusion chromatography (Superdex 75, GE Healthcare) in 20 mM Tris buffer at pH 7.1 and 150 mM NaCl. To eliminate associated sugars, the protein was dialyzed several times against 3.5 M guanidinium chloride under the same buffer conditions. Final sample conditions: 0.5 mM in 20 mM Tris, 150 mM NaCl, 10 mM CaCl$_2$, and 95%/5% H$_2$O/D$_2$O, pH 7.0.

Sample Preparation for Paramagnetic studies. Protein samples for NMR studies were prepared at a final protein concentration of 0.5 mM in 20 mM Tris (pH 7.0), 150 mM NaCl, 10 μM CaCl$_2$, and 1 mM NaN$_3$ with 10% D$_2$O. In order to conjugate the paramagnetic tag to the C-terminal domain cysteine, the protein was titrated with 10 mM solution of lanthanide (Ln) chelating tag previously loaded with the lanthanide (Ln = La^{3+}, Dy^{3+}). The titrations were performed by monitoring the changes in the chemical shift in a ^{1}H,^{15}N HSQC spectrum. The nucleophilic substitution reaction is instantaneous, and the excess of chelating tag molecule was removed by filtration.

NMR Spectroscopy. All of the NMR experiments were carried out on a 600 MHz AVANCE-III Bruker spectrometer. Spectra were acquired at 310 K. All NMR spectra were processed with the software TopSpin. The Program CARA was used for the analysis of the 2D spectra. Lanthanum was used as diamagnetic reference as it has an ionic radius similar to the paramagnetic dysprosium. PCSs were measured as the difference between the chemical shift of the corresponding nuclei in the paramagnetic and diamagnetic samples. Residual dipolar couplings ^{1}D$_{HN}$ were measured as the ^{1}H-doublet splitting of the paramagnetic sample minus the equivalent splitting difference in the diamagnetic sample.

PCS and RDC analyses were performed using MSpin software. In order to assess interdomain dynamics, we used selected conformations from molecular dynamics simulation, including some that resemble the experimentally obtained X-ray open and closed GGBP conformations. The optimal ensemble of conformations was found after evaluating the tensor from the tag domain using both PCSs and RDCs independently. The PCSs and RDCs for the other domain were back calculated for different protein coordinates, and the ensemble of structures that better fits the experimental values was selected on the basis of the best quality factor.

Molecular Dynamics Simulations. The coordinates from the X-ray structure of *apo_op* GGBP (PDB code 2FW0) and *holo_cl* GGBP (PDB code 2FVY) were used as starting points to generate intermediate models by molecular dynamics (MD) simulations at 310 K applying the ff10[33] and GLYCAM_06h[34] force fields. Missing hydrogens were added to the starting PDB structures using the program LEAP. The N- and C-terminal residues were acetylated and amidated according to the AMBER standard database. The latter structure was solvated in a cubic TIP3P water box,[35] and eight sodium ions were added to neutralize the system. In order to fill all of the proteins cavities with water molecules, a previous minimization for only solvent and ions was made. Moreover, to reach a reasonable starting structure, the entire system was minimized with a higher number of cycles, using the accurate steepest descent algorithm. The system was subjected to two rapid molecular dynamic simulations of 20 and 100 ps, respectively, before starting the real dynamic simulation of 270 ns for *apo_op* and 100 ns for *holo_cl*GGBP. During these two preparatory steps, the structure was slowly heated from 0 to 310 K. Fifty-thousand additional steps were performed to switch from constant volume to constant pressure. A relaxation time of 2 ps was used in order to equilibrate the entire system in each step. The final simulations of 270 and 100 ns were performed starting from equilibrated structures. Coordinates and energy values were recorded every 2 ps for a total of 135000 MD models for *apo_op* and 50000 for *holo_cl*GGBP. For the SMD simulations, the starting structures, with hinge values of 124° and 147° have been extracted by unrestricted MD

simulations of $holo\_cl$ and apo_opGGBP, respectively; so the entire system was already equilibrated. The center of mass of the N-terminal domain together with the center of mass of the hinge segment has been fixed, while the center of mass of the C-terminal domain has been pulled with a constant force $K = 500$ kcal·mol^{-1}·Å^2. The total time for the molecular dynamics simulation was of 10 ns with an angle opening or closing of 2° per ns. Atomic coordinates were saved every 2 ps and the energy information extracted.

Constrained MD simulations were initiated from the X-ray structures apo_op (2FW0) and $holo\_cl$ (2FVY), shown as orange and blue circles, respectively. Conjoined rigid body-torsion angle simmulated annealing was performed as previously described.[36,14] The hinge region for the GGBP is defined by residues 109−111, 253−256, and 293−296. The starting structure (apo_op, F2W0) was heated to 600 K for 3 ps with tautp equal to 0.4. Then, the system was cooled to 100 K for 297 ps (tautp = 4.0). The final cooling to 0 K was carried out for 100 ps with tautp varying from 1.0 to 0.1. The χ tensor anisotropy parameters extracted from MSpin software were those were derived from the lowest QF structures. MD simulations were performed with ff10_Amber and GLYCAM_06h force fields integrated with the experimentally derived NMR restraints. The restrained MD calculations were performed in explicit water solvent and using a simulated annealing approach.

Analysis of the Trajectories. Root mean-square deviation (simulated) and thermodynamic data were monitored throughout the whole trajectory to confirm that all simulations evolved along a stable plateau. For the analysis of the collective motions, the closure (hinge), twisting, and bending coordinate system was used.[37] Hinge and twist angles were obtained from clusters of one in every 100 models (1350 in total) for apo_op and one in every 50 models for the closed bound GGBP conformation ($holo\_cl$GGBP). For all structures, the values of hinge θ and twist ϕ angles were calculated using an in-house program in Matlab to adequately represent the conformational landscape in terms of the hinge and twist angles. For the evaluation of the tag-domain data, the structures were aligned with respect to the backbone of residues 112−254. The tag molecule was accommodated for all of the structures, and a rapid minimization on the tag region was made. The structures were visualized and evaluated by using the programs VMD and Discovery Studio.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschem-bio.6b00148.

> Synthesis and chemical characterization of the tag molecule; description of the MSpin calculations; and detailed tables with the PCSs and RDCs measured with the two metal ions (La^{3+} and Dy^{3+}) and parameters of the calculated $\Delta\chi$ tensor (PDF)

> Movie illustrating the dynamic behavior of GGBP free in solution and in the presence of the ligand (MP4)

AUTHOR INFORMATION

Corresponding Authors

*(J.J.-B.) E-mail: jjbarbero@cicbiogune.es.

*(O.M.) E-mail: omillet@cicbiogune.es.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Shilton, B. H., Flocco, M. M., Nilsson, M., and Mowbray, S. L. (1996) Conformational changes of three periplasmic receptors for bacterial chemotaxis and transport: the maltose-, glucose/galactose-and ribose-binding proteins. *J. Mol. Biol.* 264, 350−363.

(2) Tam, R., and Saier, M. H., Jr. (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* 57, 320−346.

(3) Loeffler, H. H., and Kitao, A. (2009) Collective dynamics of periplasmic glutamine binding protein upon domain closure. *Biophys. J.* 97, 2541−2549.

(4) Stockner, T., Vogel, H. J., and Tieleman, D. P. (2005) A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. *Biophys. J.* 89, 3362−3371.

(5) Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1991) Comparison of the periplasmic receptors for L-arabinose, D-glucose/D-galactose, and D-ribose. Structural and Functional Similarity. *J. Biol. Chem.* 266, 5226−5237.

(6) Borrok, M. J., Kiessling, L. L., and Forest, K. T. (2007) Conformational changes of glucose/galactose-binding protein illuminated by open, unliganded, and ultra-high-resolution ligand-bound structures. *Protein Sci.* 16, 1032−1041.

(7) Borrok, M. J., Zhu, Y., Forest, K. T., and Kiessling, L. L. (2009) Structure-based design of a periplasmic binding protein antagonist that prevents domain closure. *ACS Chem. Biol.* 4, 447−456.

(8) Cuneo, M. J., Changela, A., Warren, J. J., Beese, L. S., and Hellinga, H. W. (2006) (2006) The crystal structure of a thermophilic glucose binding protein reveals adaptations that interconvert mono and di-saccharide binding sites. *J. Mol. Biol.* 362, 259−270.

(9) Sooriyaarachchi, S., Ubhayasekera, W., Boos, W., and Mowbray, S. L. (2009) X-ray structure of glucose/galactose receptor from Salmonella typhimurium in complex with the physiological ligand, (2R)-glyceryl-beta-D-galactopyranoside. *FEBS J.* 276, 2116−2124.

(10) Tian, Y., Cuneo, M. J., Changela, A., Hocker, B., Beese, L. S., and Hellinga, H. W. (2007) Structure-based design of robust glucose biosensors using a Thermotoga maritima periplasmic glucose-binding protein. *Protein Sci.* 16, 2240−2250.

(11) Millet, O., Hudson, R. P., and Kay, L. E. (2003) The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12700−12705.

(12) Tang, C., Schwieters, C. D., and Clore, G. M. (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* 449, 1078−1082.

(13) Bermejo, G. A., Strub, M. P., Ho, C., and Tjandra, N. (2010) Ligand-free open-closed transitions of periplasmic binding proteins: the case of glutamine-binding protein. *Biochemistry* 49, 1893−1902.

(14) Ortega, G., Castano, D., Diercks, T., and Millet, O. (2012) Carbohydrate affinity for the glucose-galactose binding protein is regulated by allosteric domain motions. *J. Am. Chem. Soc.* 134, 19869−19876.

(15) Ravindranathan, K. P., Gallicchio, E., and Levy, R. M. (2005) Conformational equilibria and free energy profiles for the allosteric transition of the ribose-binding protein. *J. Mol. Biol.* 353, 196−210.

(16) Wereszczynski, J., and McCammon, J. A. (2012) Statistical mechanics and molecular dynamics in evaluating thermodynamic properties of biomolecular recognition. *Q. Rev. Biophys.* 45, 1−25.

(17) Dwyer, M. A., and Hellinga, H. W. (2004) Periplasmic binding proteins: a versatile superfamily for protein engineering. *Curr. Opin. Struct. Biol.* 14, 495−504.

(18) Jeffery, C. J. (2011) Engineering periplasmic ligand binding proteins as glucose nanosensors. *Nano Rev.* 2, 5743.

(19) Vercillo, N. C., Herald, K. J., Fox, J. M., Der, B. S., and Dattelbaum, J. D. (2007) Analysis of ligand binding to a ribose biosensor using site-directed mutagenesis and fluorescence spectroscopy. *Protein Sci.* 16, 362−368.

(20) Salins, L. L., Ware, R. A., Ensor, C. M., and Daunert, S. (2001) A novel reagentless sensing system for measuring glucose based on the galactose/glucose-binding protein. *Anal. Biochem. 294*, 19−26.

(21) Russo, L., Maestre-Martinez, M., Wolff, S., Becker, S., and Griesinger, C. (2013) Interdomain dynamics explored by paramagnetic NMR. *J. Am. Chem. Soc. 135*, 17111−17120.

(22) Canales, A., Mallagaray, A., Perez-Castells, J., Boos, I., Unverzagt, C., Andre, S., Gabius, H. J., Canada, F. J., and Jimenez-Barbero, J. (2013) Breaking pseudo-symmetry in multiantennary complex N-glycans using lanthanide-binding tags and NMR pseudo-contact shifts. *Angew. Chem., Int. Ed. 52*, 13789−13793.

(23) Bertini, I., Del Bianco, C., Gelis, I., Katsaros, N., Luchinat, C., Parigi, G., Peana, M., Provenzani, A., and Zoroddu, M. A. (2004) Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. *Proc. Natl. Acad. Sci. U. S. A. 101*, 6841−6846.

(24) Haberz, P., Rodriguez-Castaneda, F., Junker, J., Becker, S., Leonov, A., and Griesinger, C. (2006) Two new chiral EDTA-based metal chelates for weak alignment of proteins in solution. *Org. Lett. 8*, 1275−1278.

(25) Tjandra, N., Kuboniwa, H., Ren, H., and Bax, A. (1995) Rotational dynamics of calcium-free calmodulin studied by 15N-NMR relaxation measurements. *Eur. J. Biochem. 230*, 1014−1024.

(26) Gao, M., Wilmanns, M., and Schulten, K. (2002) Steered molecular dynamics studies of titin I1 domain unfolding. *Biophys. J. 83*, 3435−3445.

(27) Krammer, A., Lu, H., Isralewitz, B., Schulten, K., and Vogel, V. (1999) Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc. Natl. Acad. Sci. U. S. A. 96*, 1351−1356.

(28) Gu, J., Li, H., and Wang, X. (2015) A Self-Adaptive Steered Molecular Dynamics Method Based on Minimization of Stretching Force Reveals the Binding Affinity of Protein-Ligand Complexes. *Molecules 20*, 19236−19251.

(29) Harris, S. A., Sands, Z. A., and Laughton, C. A. (2005) Molecular dynamics simulations of duplex stretching reveal the importance of entropy in determining the biomechanical properties of DNA. *Biophys. J. 88*, 1684−1691.

(30) Perisic, O., and Lu, H. (2014) On the improvement of free-energy calculation from steered molecular dynamics simulations using adaptive stochastic perturbation protocols. *PLoS One 9*, e101810.

(31) Schubert, H. L., Phillips, J. D., Heroux, A., and Hill, C. P. (2008) Structure and mechanistic implications of a uroporphyrinogen III synthase-product complex. *Biochemistry 47*, 8648−8655.

(32) Flocco, M. M., and Mowbray, S. L. (1994) The 1.9 A x-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from Salmonella typhimurium. *J. Biol. Chem. 12*, 8931−8936.

(33) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet. 65*, 712−725.

(34) Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem. 29*, 622−55.

(35) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.

(36) Clore, G. M., and Bewley, C. A. (2002) Using Conjoined Rigid Body/Torsion Angle Simulated Annealing to Determine the Relative Orientation of Covalently Linked Protein Domains from Dipolar Couplings. *J. Magn. Reson. 154*, 329−335.

(37) Skrynnikov, N. R., Goto, N. K., Yang, D., Choy, W. Y., Tolman, J. R., Mueller, G. A., and Kay, L. E. (2000) Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin. *J. Mol. Biol. 295*, 1265−1273.