

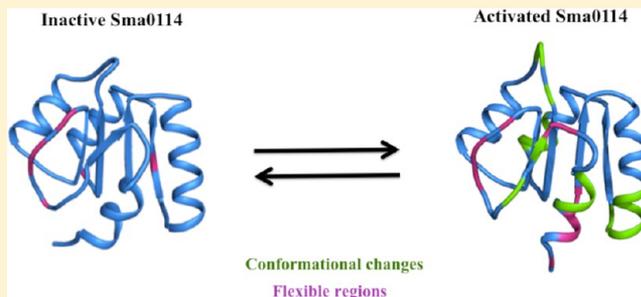
# NMR Structure of the HWE Kinase Associated Response Regulator Sma0114 in Its Activated State

Sarah R. Sheftic, Emma White, Daniel J. Gage, and Andrei T. Alexandrescu\*

Department of Molecular and Cellular Biology, University of Connecticut, Storrs, Connecticut, United States

## S Supporting Information

**ABSTRACT:** Bacterial receiver domains modulate intracellular responses to external stimuli in two-component systems. Sma0114 is the first structurally characterized representative from the family of receiver domains that are substrates for histidine-tryptophan-glutamate (HWE) kinases. We report the NMR structure of Sma0114 bound by  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$ , a phosphate analogue that stabilizes the activated state. Differences between the NMR structures of the inactive and activated states occur in helix  $\alpha 1$ , the active site loop that connects strand  $\beta 3$  and helix  $\alpha 3$ , and in the segment from strand  $\beta 5$  to helix  $\alpha 5$  of the 455 ( $\alpha 4$ - $\beta 5$ - $\alpha 5$ ) face. Structural rearrangements of the 455 face typically make receiver domains competent for binding downstream target molecules. In Sma0114 the structural changes accompanying activation result in a more negatively charged surface for the 455 face. Coupling between the 455 face and active site phosphorylation is usually mediated through the rearrangement of a threonine and tyrosine residue, in a mechanism called Y–T coupling. The NMR structure indicates that Sma0114 lacks Y–T coupling and that communication between the active site and the 455 face is achieved through a conserved lysine residue that stabilizes the acyl phosphate in receiver domains.  $^{15}\text{N}$ -NMR relaxation experiments were used to investigate the backbone dynamics of the Sma0114 apoprotein, the binary Sma0114· $\text{Ca}^{2+}$  complex, and the ternary Sma0114· $\text{Ca}^{2+}$ · $\text{BeF}_3^-$  complex. The loss of entropy due to ligand binding at the active site is compensated by increased flexibility in the 455 face. The dynamic character of the 455 face in Sma0114, which results in part from the replacement of helix  $\alpha 4$  by a flexible loop, may facilitate induced-fit recognition of target molecules.



Phosphorylation-mediated signal transduction in bacteria is carried out by “two-component systems”.<sup>1,2</sup> The first component is a membrane-bound sensor histidine kinase that autophosphorylates when triggered by an external stimulus. The second component is a response regulator that mediates the output response, after becoming phosphorylated by its cognate histidine kinase. Many response regulators are multidomain proteins that have receiver and effector domains.<sup>1,2</sup> A smaller subset of about 15% of response regulators are single domain proteins, containing only the receiver domain.<sup>3–5</sup> Of the ~30 000 nonredundant receiver domain sequences known, ~200 have been structurally characterized.<sup>3</sup> Receiver domain structures have a conserved  $\alpha_5/\beta_5$  Rossmann-fold topology with a five-stranded parallel  $\beta$ -sheet surrounded by five  $\alpha$ -helices.<sup>1–3</sup> Activation of receiver domains is accomplished by phosphorylation at a conserved aspartate in the presence of a divalent metal ion such as  $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$ , or  $\text{Ca}^{2+}$ .<sup>6</sup> Upon phosphorylation, receiver domains undergo a conformational transition to an activated state, becoming primed for binding the downstream effector molecules that elicit the signal transduction output response. Receiver domains have autophosphatase enzymatic activity that controls the lifetime of the activated state and hence the duration of the output response.<sup>7,8</sup>

The gene *sma0114* of *Sinorhizobium meliloti* encodes a single domain response regulator that is a substrate for a histidine-tryptophan-glutamate (HWE) kinase.<sup>9,10</sup> Sma0114 is the only representative of the HWE kinase-associated receiver domain family with a known structure.<sup>11</sup> Bioinformatic and structural analysis of Sma0114 revealed differences from the canonical receiver domain family.<sup>11</sup> In the inactive enzyme, a disordered region replaces the prototypical helix  $\alpha 4$  of the 455 ( $\alpha 4$ - $\beta 5$ - $\alpha 5$ ) face, which usually undergoes the largest structural changes on activation to form a binding interface for downstream effector molecules. In typical receiver domains, the phosphorylation state of the active site is communicated to the 455 face through a conformational rearrangement called “Y–T coupling”, involving a structurally conserved tyrosine (Y) and threonine (T).<sup>2</sup> In Sma0114, the tyrosine of Y–T coupling is replaced by a leucine. This suggests that Sma0114 lacks a Y–T coupling mechanism, as has been previously reported for the receiver domain CheY2.<sup>12</sup> Sma0114 has a (PFx FATGY) sequence motif that is highly conserved in the subfamily of receiver domains associated with HWE kinases.<sup>11</sup> On the basis of the NMR

Received: November 5, 2013

Revised: December 22, 2013

Published: December 23, 2013



structure of the inactive state, we suggested that the PFx FATGY motif might substitute for the typical Y–T coupling mechanism in Sma0114.

To better understand the mechanism of activation, we determined the NMR structure of Sma0114 in its activated state. Because the lifetimes of acyl phosphates (~30 min) are too short to allow detailed NMR studies, we used the phosphate mimetic beryllium trifluoride ( $\text{BeF}_3^-$ ) to stabilize the activated state, as has been done for other structural studies of receiver domains.<sup>13,14</sup> The structures of inactive and activated Sma0114 were analyzed for key differences that could communicate the phosphorylation state of the receiver domain from the active site to the 455 face of the protein. We compared the conformational changes accompanying activation of Sma0114 to three homologous single-domain receiver domains, to see what features are conserved. Complementary  $^{15}\text{N}$  relaxation experiments were used to investigate how the dynamics of Sma0114 are affected by binding  $\text{Ca}^{2+}$  and by activation with  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$ .

## EXPERIMENTAL PROCEDURES

**NMR Sample Preparation.** The expression and purification of Sma0114 were performed as described previously.<sup>15</sup> For NMR spectroscopy,  $^{15}\text{N}$ -Sma0114 and  $^{15}\text{N},^{13}\text{C}$ -Sma0114 samples were dissolved in 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES), buffered to pH 6.0. Samples had 500  $\mu\text{M}$  concentrations of protein, with 0.02%  $\text{NaN}_3$  to prevent bacterial growth and 1 mM dithiothreitol (DTT) to prevent disulfide formation due to the sole cysteine at position 29 in the protein. All NMR experiments were done at a temperature of 37 °C. For work on the binary Sma0114- $\text{Ca}^{2+}$  complex, we used a 1.5 mM concentration of  $\text{CaCl}_2$ , which we previously determined by NMR is sufficient to saturate the active site.<sup>16</sup> The Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex was prepared using the approach of Yan et al.<sup>13</sup>  $\text{BeF}_3^-$ , generated in situ from the addition of NaF to  $\text{BeCl}_2$ , mimics the acyl phosphate linkages formed by the active site aspartates of receiver domains.<sup>13,14,17,18</sup> Starting from samples of Sma0114 containing 1.5 mM  $\text{CaCl}_2$ , we added NaF in increments of 5 mM and observed no changes in  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the protein up to a final concentration of 30 mM NaF. Thus, the presence of NaF, which is needed to generate  $\text{BeF}_3^-$  from  $\text{BeCl}_2$ , does not perturb the structure of Sma0114. We next added  $\text{BeCl}_2$  in 1 mM increments and found that a 5 mM concentration was sufficient to saturate the active site, as monitored by the point at which the peaks from the apo-state, which in slow exchange with the activated state, disappeared, and there were no further chemical shift changes in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the activated state. Thus, the final sample conditions for the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex were 500  $\mu\text{M}$  Sma0114, 1.5 mM  $\text{CaCl}_2$ , 30 mM NaF, and 5 mM  $\text{BeCl}_2$ .

**NMR Structure Determination.** Chemical shifts for the Sma0114- $\text{Ca}^{2+}$  binary complex were obtained from the titration of  $^1\text{H}$ - $^{15}\text{N}$  crosspeaks in 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Sma0114 as a function of  $\text{CaCl}_2$  concentration.<sup>11</sup> NMR assignments for the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex were made from 2D and 3D NMR experiments as previously described for the apoprotein.<sup>15</sup> The chemical shift assignments for the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex have been deposited in the BMRB database under accession code 19286. Distance restraints for structure calculations were obtained from 3D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOESY experiments<sup>19</sup> collected on a Varian Inova 800 MHz instrument, and a 2D  $^1\text{H}$ -NOESY experiment for the

aromatic protons recorded at 600 MHz. The mixing time for all NOESY experiments was 150 ms. Hydrogen bonds were identified based on solvent protection in deuterium isotope exchange experiments recorded on the ternary complex dissolved in 99.96%  $\text{D}_2\text{O}$  and NOESY data. The protection pattern for the activated state was the same as that of the inactivated enzyme whose structure we determined previously,<sup>11</sup> indicating no large differences in backbone hydrogen bonding. Backbone  $\phi$  and  $\psi$  dihedral angles were calculated from the assigned HN, H $\alpha$ , N, C $\alpha$ , C $\beta$ , and C' chemical shifts using the program DANGLE in CCPNMR analysis.<sup>20</sup>

The NMR structure of the activated Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex was calculated using the program X-PLOR (v. 3.851)<sup>21</sup> based on 1791 experimental restraints (Table 1). Calculations were performed using  $\text{Ca}^{2+}$ , Be, and F atomic radii from the X-plor library<sup>21</sup> to include the divalent metal ion and  $\text{BeF}_3^-$ . Because there are no literature examples available for a bond between  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$ , we set the distance restraint

**Table 1. Statistics for the 26 Lowest Energy NMR Structures of Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$**

Experimental Restraints		
NMR restraints (total)		1791
distance (total)		1644
intraresidue NOEs		472
sequential NOEs		479
short range NOEs ( $1 <  i - j  < 5$ )		136
long range NOEs ( $5 \leq  i - j $ )		495
hydrogen bonds ( $31 \times 2$ )		62
dihedral ( $\phi$ 70, $\psi$ 61, $\chi_1$ 16)		147
Residual Restraint Violations		
NOE (Å) <sup>a</sup>		0.033 ± 0.0004
dihedral (°) <sup>b</sup>		0.587 ± 0.075
RMSD from ideal geometry		
bonds (Å)		0.0037 ± 0.0002
angles (°)		0.724 ± 0.004
improper torsions (°)		0.560 ± 0.007
van der Waals energy (kcal/mol) <sup>c</sup>		87.85 ± 1.75
Lennard–Jones energy (kcal/mol) <sup>d</sup>		−85.42 ± 6.33
ProcheckNMR Z-score <sup>e</sup>		−2.11
Ramachandran statistics for all residues <sup>e</sup> (and excluding loops <sup>f</sup> )		
most favored (%)	58.8	(58.2)
allowed (%)	29.6	(36.8)
generously allowed (%)	7.5	(4.9)
disallowed (%)	4.1	(0.0)
	Coordinate RMSD (Å)	
NMR ensemble to average	<u>C<math>\alpha</math></u> , C, N	all heavy
entire domain (110 residues) <sup>g</sup>	0.88	1.34
excluding loops (70 residues) <sup>f</sup>	0.71	1.19
NMR model to inactive model		
entire domain (110 residues) <sup>g</sup>	1.91	2.70
excluding loops (70 residues) <sup>f</sup>	1.66	2.32

<sup>a</sup>Structures have no NOE violations greater than 0.3 Å. <sup>b</sup>Structures have no dihedral violations greater than 5°. <sup>c</sup> $E_{\text{vdw}}$  was calculated using the X-PLOR Frepel function<sup>21</sup> with van der Waals interactions and atomic radii set to 0.8 times their CHARMM<sup>42</sup> values. <sup>d</sup> $E_{\text{LJ}}$  was calculated using the CHARMM empirical energy function.<sup>42</sup> <sup>e</sup>ProcheckNMR used via Protein Structure Validation Suite (PSVS).<sup>43</sup> <sup>f</sup>Only residues in regular secondary structure (10–15, 18–31, 34–39, 42–51, 55–60, 70–78, 81–86, 102–104, 110–118). <sup>g</sup>Excluding N-terminus (residues 1–9) and C-terminus (residues 119–123).

between the two atoms at 2.35 Å, equal to the sum of the ionic radii for  $\text{Ca}^{2+}$  (0.99 Å) and  $\text{F}^-$  (1.36 Å). The distance restraint was given an uncertainty bound of 0.1 Å, and the three fluorine atoms in  $\text{BeF}_3^-$  were treated as ambiguous. We next included an ambiguous restraint of  $1.5 \pm 0.3$  Å between the two side chain carboxylic acid oxygens of Asp60 (OD1 or OD2) and the beryllium atom. The restraint was based on distances between the beryllium atom and the closest oxygen of the equivalent aspartate in representative X-ray structures of receiver domains (PDB codes: 1FWQ, 2A9O, 2PL1, 3NNN). Bond distances in these structures ranged from 1.5 to 2.0 Å, with an average distance of 1.7 Å. We used the same approach to restrain the three possible  $\text{Ca}^{2+}$  ligands in the protein, Glu15-Asp16-Glu17. An analogous stretch of two or three acidic residues in a row is found in the receiver domain structures listed above. To determine which of the three acidic residues in Sma0114 could be ligands for  $\text{Ca}^{2+}$ , we carried out structure calculations with the distance restraint set for the complete structure but excluding restraints to  $\text{Ca}^{2+}$  or  $\text{BeF}_3^-$ . Superposition of the two structures showed that both Glu15 and Asp16 were within bonding distance of the  $\text{Ca}^{2+}$  in the structures calculated without restraints to the ligands. The two glutamates are also within bonding distance to the  $\text{Ca}^{2+}$ , when the structure with restraints to the ligands was superposed with the NMR structure of inactive Sma0114.<sup>11</sup> That is, the two residues already appear to be positioned to bind the metal in the inactive apo-state. Glu17 points away from the  $\text{Ca}^{2+}$ , but the conformation of this residue is poorly defined due to R2 line broadening. We therefore decided to also include Glu17 as a potential  $\text{Ca}^{2+}$  binding ligand. Distance restraints were set to  $2.45 \pm 0.15$  Å based on literature values of typical bonds between carboxylic acid groups and  $\text{Ca}^{2+}$  in known protein structures,<sup>22</sup> which range between 2.3 and 2.6 Å. The distance restraints were treated as ambiguous with respect to the oxygens bound to the  $\text{Ca}^{2+}$ : Glu15 (OE1 or OE2), Asp16 (OD1 or OD2), and Glu17 (OE1 or OE2). Finally, we introduced a restraint of  $2.5 \pm 0.5$  Å between the  $\text{N}_\alpha$  atom of Lys105 and the fluorine atoms of  $\text{BeF}_3^-$ , which were treated as ambiguous in the calculations. An analogous lysine is found bound to  $\text{BeF}_3^-$  in all  $\text{BeF}_3^-$  activated receiver domain structures, with a bonding distance between 2.8 and 3.0 Å. To test the effect of this restraint, we calculated a set of structures in which only the Lys105- $\text{BeF}_3^-$  restraint was removed. Without the restraint, the lysine side chain was positioned in the interior of the structure, but the distance of 6.0 Å to  $\text{BeF}_3^-$  precluded a bonding interaction. Given that a buried charged lysine would be highly energetically unfavorable<sup>23</sup> as there are no other nearby negatively charged groups that could complement the positive charge on Lys105, we felt the restraint to the  $\text{BeF}_3^-$  ion was justified. The restraints from the protein to the  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$  ligands described above, introduced no violations in the structure calculations. The RMSD between the structures calculated with and without restraints to the ligands was 0.80 Å, comparable to the 0.88 Å value for the spread in the ensemble of activated state NMR structures (Table 1). This indicates that the derived restraints to the ligands do not affect the overall structure and that the structure with the restraints to the ligands is compatible with the experimental NOE, hydrogen-bond, and dihedral restraints. The 26 lowest energy structures for the activated state of Sma0114 have been deposited in the PDB under accession code 2m98.

**NMR Relaxation Measurements.** All NMR relaxation experiments were done at a field strength of 600 MHz. Backbone dynamics of the Sma0114- $\text{Ca}^{2+}$  binary and Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complexes were characterized using  $^{15}\text{N}$   $R_1$ ,  $R_2$ , and  $^1\text{H}$ - $^{15}\text{N}$  NOE experiments, shown in Supplementary Figures S1 and S2, Supporting Information. Longitudinal relaxation rates ( $R_1$ ) were determined using relaxation delays of 0.02, 0.05, 0.13, 0.21, 0.31, 0.5, 0.71, and 1.0 s. Transverse relaxation rates ( $R_2$ ) were determined using relaxation delays of 0.01, 0.03, 0.05, 0.07, 0.09, 0.15, 0.25, and 0.35 s.  $^1\text{H}$ - $^{15}\text{N}$  NOE values were measured from experiments in which the proton spectrum was saturated for 4 s and control experiments in which the saturation period was replaced by an equivalent preacquisition delay. Relaxation rates were calculated from least-squares fits of the data to an exponential decay model (eq 1), where  $I$  is the intensity for the relaxation period  $\tau$ ,  $I_0$  is the initial amplitude, and  $R_{1,2}$  corresponds to the relaxation rate  $R_1$  or  $R_2$ .

$$I = I_0 X \exp(-\tau/R_{1,2}) \quad (1)$$

$^1\text{H}$ - $^{15}\text{N}$  NOE values were calculated according to eq 2, where  $I(s)$  is the crosspeak intensity in the experiment with saturation and  $I(c)$  is the crosspeak intensity without saturation.

$$\text{NOE} = I(s)/I(c) \quad (2)$$

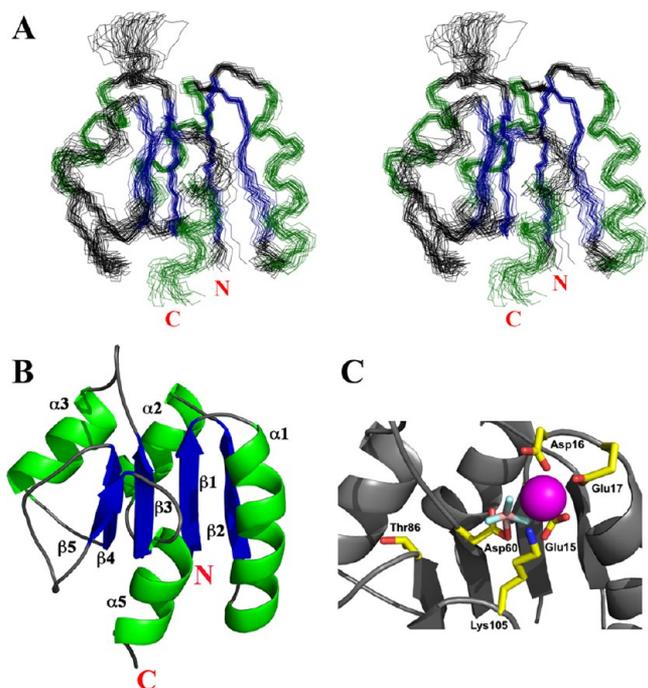
Experimental uncertainties for the relaxation parameters were determined as described previously.<sup>24,25</sup> Model Free calculations<sup>26</sup> using the program *Tensor 2*<sup>27</sup> were performed to characterize the dynamics of the binary and ternary complexes in terms of amplitudes of motions ( $S^2$  order parameters) and chemical exchange contributions to transverse relaxation rates ( $R_{2\text{ex}}$ ).

## RESULTS

**NMR Structure of Activated Sma0114.** Backbone ( $\text{C}^\alpha$ ,  $\text{N}$ ,  $\text{C}'$ ) traces for the 26 lowest energy NMR structures of the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex are shown in Figure 1A. Statistics pertaining to the quality of the NMR structures are given in Table 1. The precision of the activated NMR structure over elements of regular secondary structure is 0.71 Å, slightly better than the 0.86 Å precision of the inactive state.<sup>11</sup> Other measures of the quality of the structure including the Procheck Z-score, RMS deviations from ideal geometry, and average residual NOE violations are also slightly better for the activated than inactive structure.<sup>11</sup>

A cartoon showing the  $\alpha_5/\beta_5$  Rossmann fold topology of the activated structure is shown in Figure 1B. As previously described for the structure of the inactive state,<sup>11</sup> Sma0114 differs from typical receiver domains<sup>2</sup> in having the fourth  $\alpha$ -helix of the 455 face replaced by a flexible loop. The fourth helix is also absent in the structure of the activated state; thus binding of  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$  does not induce folding of this segment.

Figure 1C shows the active site and key residues that interact with the  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$  ions. The three residues Glu15, Asp16, and Glu17 in the loop between strand  $\beta_1$  and helix  $\alpha_1$  ( $L_{\beta_1\alpha_1}$ ) each have one oxygen within bonding distance of the  $\text{Ca}^{2+}$  ion (2.1, 2.1, and 2.4 Å). The  $\text{Ca}^{2+}$  ion is within bonding distance of two fluorine atoms from the  $\text{BeF}_3^-$  molecule (1.8 and 2.1 Å). The third fluorine atom does not appear to form any bonding interactions. The fluorine atom closest to the  $\text{Ca}^{2+}$  is within a 2.6 Å bonding distance of the  $\text{N}_\alpha$  atom of Lys105. The

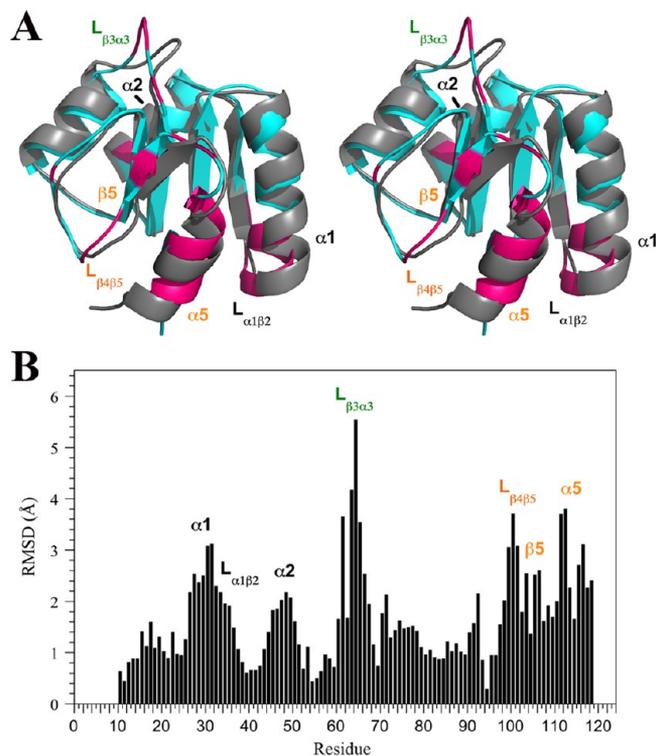


**Figure 1.** NMR Structure of the ternary Sma0114-Ca<sup>2+</sup>·BeF<sub>3</sub><sup>-</sup> complex. (A) Stereodiamgram of the 26 lowest energy NMR structures superimposed over regular main-chain secondary structure elements:  $\beta 1$  (10–15),  $\alpha 1$  (18–31),  $\beta 2$  (34–39),  $\alpha 2$  (42–51),  $\beta 3$  (55–60),  $\alpha 3$  (70–78),  $\beta 4$  (81–86),  $\beta 5$  (102–104),  $\alpha 5$  (110–118). (B) Backbone cartoon of the NMR structure closest to the ensemble mean. (C) Sidechains that participate in ligand binding at the active site of Sma0114 and Thr86. The Ca<sup>2+</sup> ion is shown as a purple sphere and the BeF<sub>3</sub><sup>-</sup> molecule with pink and cyan sticks.

positively charged N<sub>z</sub> atom of Lys105 may also be involved in an electrostatic interaction with the OE2 atom of Glu15, with the two separated by 3.9 Å. The Be atom of BeF<sub>3</sub><sup>-</sup> is bonded to the OD2 atom of Asp60 (2.0 Å).

**Conformational Changes Induced by the Activation of Sma0114.** The structures of Sma0114 in the inactive and in the activated state, modeled by the Sma0114-Ca<sup>2+</sup>·BeF<sub>3</sub><sup>-</sup> ternary complex, are shown superimposed in Figure 2A. A plot of the C<sub>α</sub> RMSD after superposition of the two structures is shown in Figure 2B. The mean RMSD over all residues is 1.91 Å (Table 1). Regions of the protein that show larger differences than the mean RMSD are colored in pink and labeled (Figure 2A).

The largest difference is seen for the loop between strand  $\beta 3$  and helix  $\alpha 3$  (L <sub>$\beta 3\alpha 3$</sub> ) that partially covers the active site in the inactive structure but is in an open and extended conformation in the activated state. As expected, differences are seen in the 455 face. The C-terminus of strand  $\beta 5$  becomes shorter by one residue, Lys105, which becomes part of the L <sub>$\beta 5\alpha 5$</sub>  loop in the activated state. The C-terminus of the short strand  $\beta 5$  moves away from the core of the structure by about 1.0–1.5 Å in the activated state and undergoes a twist of about 30 deg to align better with strand  $\beta 4$ . In the activated state, helix  $\alpha 5$  undergoes translational displacements in the activated state, including a shift of 1.0 Å toward the top of the molecule (defined by the arrows indicating the C-termini of the  $\beta$ -strands in Figure 2A) and a move of about 1.5 Å closer to the core of structure. Outside of the 455 face, there are changes in  $\alpha$ -helices 1 and 2 (Figure 2). Helix  $\alpha 1$  undergoes a bend in the activated structure which brings the last two turns about 1–2 Å closer to helix  $\alpha 5$

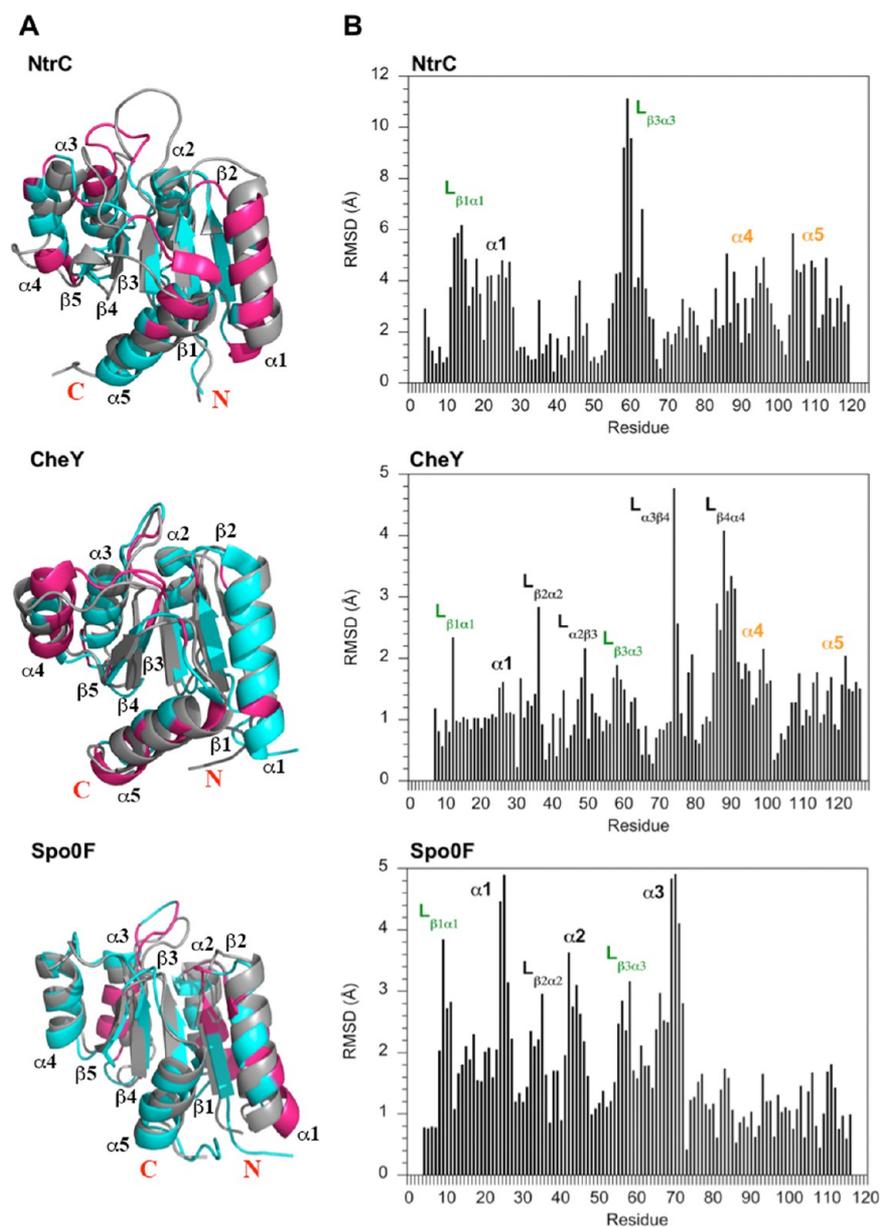


**Figure 2.** Comparison of the inactive and activated NMR structures of Sma0114. (A) Stereodiamgram showing the least-squares superposition of the inactive (gray) and activated (cyan) structures. Residues with the largest RMS differences between the two structures are indicated in pink on the activated structure. (B) Bar plot of C<sub>α</sub> RMSD values between the inactive and activated structures. Labels indicate the largest differences between the two structures, with green and orange labels indicating segments from the active site and 455 face, respectively.

and to the core of the structure. Similarly, helix  $\alpha 2$  experiences a 10–20 deg axis tilt in the activated state, which brings the C-terminus of the  $\alpha$ -helix closer to the  $\beta$ -sheet in the core of the protein. Overall, the movement of  $\alpha$ -helices 1 and 5 toward each other and the protein core make the activated state structure slightly more compact than the inactivated structure. The radii of gyration calculated from the structures are 14.2 Å for the inactive and 13.8 for the activated states.

**Comparison to Other Receiver Domains.** To investigate the extent to which the changes associated with activation of Sma0114 are conserved in other receiver domains, we looked at three single-domain response regulators NtrC, CheY, and Spo0F, shown in Figure 3. As in Figure 2, the largest changes between the inactive and activated structures are highlighted in pink on the activated structures (Figure 3A), and bar graphs are used to summarize C<sub>α</sub> RMSDs (Figure 3B). The analysis shows that the conformational changes induced by the activation of receiver domains can be quite variable.

In NtrC and CheY, conformational changes occur primarily along the 455 face of the activated enzymes, particularly for helices  $\alpha 4$  and  $\alpha 5$ . Despite the absence of helix  $\alpha 4$  in Sma0114, we do observe changes in the structure of the  $\beta 5$ - $\alpha 5$  segment (Figure 2). The Spo0F receiver domain shows only small differences for the 455 face, with the majority of structural changes occurring for the first three  $\alpha$ -helices outside of the 455 face (Figure 3).<sup>17,28</sup> Like other single domain response regulators (Figure 3), Sma0114 shows structural changes for

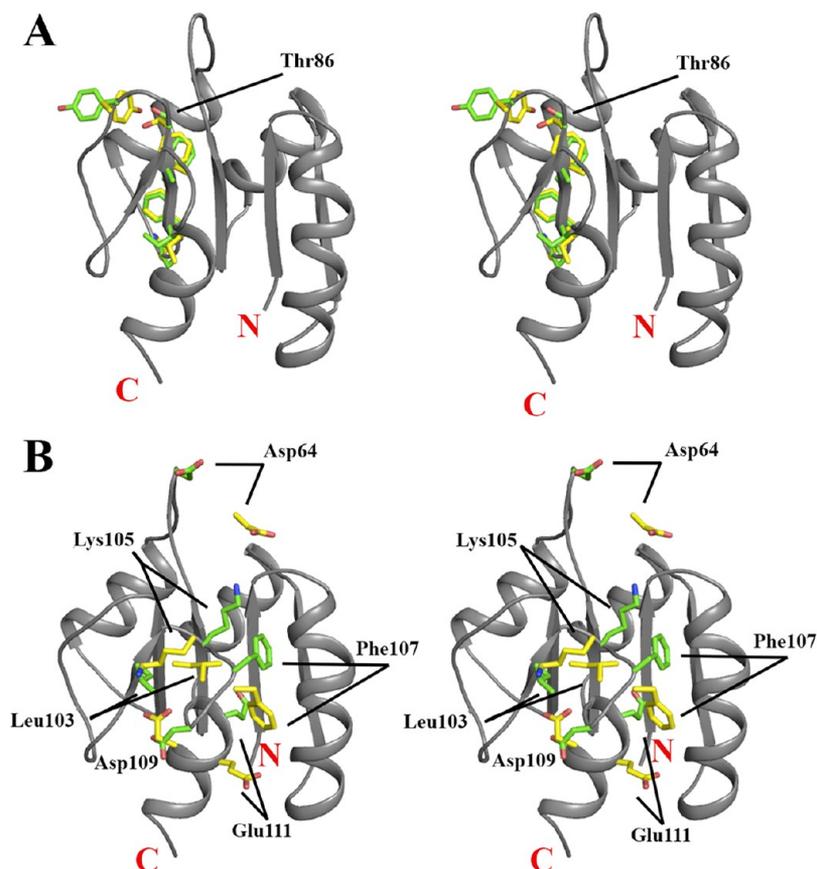


**Figure 3.** Comparison of inactive and activated states in structurally homologous single domain response regulators. (A) Superposition of inactive (gray) and activated (cyan) structures of NtrC (PDB codes 1NTR and 1KRX), CheY (PDB codes 5CHT and 1DJM), and Spo0F (PDB codes 1NAT and 1PUX). Regions that show the largest differences between the structures are colored in pink on the activated structures. (B) Corresponding  $C\alpha$  RMSD plots showing the largest differences between the inactive and activated structures. Green and orange labels indicate segments from the active site and 455 face, respectively. Residues from the unstructured N- and C-termini of all three proteins are excluded from the plots.

helix  $\alpha 1$  between the inactive and active states (Figure 2). This suggests that helix  $\alpha 1$  could participate together with the 455 face in binding downstream effectors, or alternatively that it may be important in the transition to an activated state. Indeed, mutagenesis experiments indicate that helix  $\alpha 1$  plays a role in determining the kinase specificity of receiver domains.<sup>29</sup> The types of structural changes that occur upon activation appear to follow similar trends. As in Sma0114, activation of the NtrC, CheY, and Spo0F single domain receiver domains brings the  $\alpha$ -helices closer to the  $\beta$ -sheet core of the structures (Figure 3A). Another conserved theme is that helices  $\alpha 1$  and  $\alpha 5$  move closer together in the activated states (Figure 3A).

The largest changes observed between the inactive and activated states are in the active site loops, particularly  $L_{\beta 3\alpha 3}$

following the phosphorylation site (Figure 3B). In all three structures, except CheY, where the differences are relatively small,  $L_{\beta 3\alpha 3}$  moves away from the active site in the activated state (Figure 3A). The open loop conformation of  $L_{\beta 3\alpha 3}$  was noted for the activated structure of NtrC,<sup>18</sup> but the authors noted that the precision of that region in the inactive structure precluded additional analysis. In Sma0114, the differences between the conformations of residues in  $L_{\beta 3\alpha 3}$  from the inactive and activated state (average RMSD of 3.2 Å for residues 62–67) are much larger than the internal precisions of the structures (RMSDs of 0.94 for the inactive<sup>11</sup> and 0.88 for the activated structure (Table 1). That the displacement of  $L_{\beta 3\alpha 3}$  appears to be a conserved theme for the activated states of



**Figure 4.** Stereodiagrams comparing selected sidechains between the inactive and activated states of Sma0114. (A) The sidechain conformations of the PFXFATGY motif are conserved between the inactive and activated state. The PFXFATGY motif contains Thr86, which is in a homologous position to the threonine of the classical receiver domain Y–T coupling mechanism. (B) Sidechains that undergo large structural changes between the inactive and activated conformations include Asp64 at the active site, and Leu103, Lys105, Phe107, and Glu111 in the 455 face of the domain. The backbone trace structure is shown for the activated state (gray). Sidechains from the inactive structure are shown with yellow carbons, while those from the activated state are shown with green carbons.

receiver domains suggests that its open conformation plays an important role in activation.

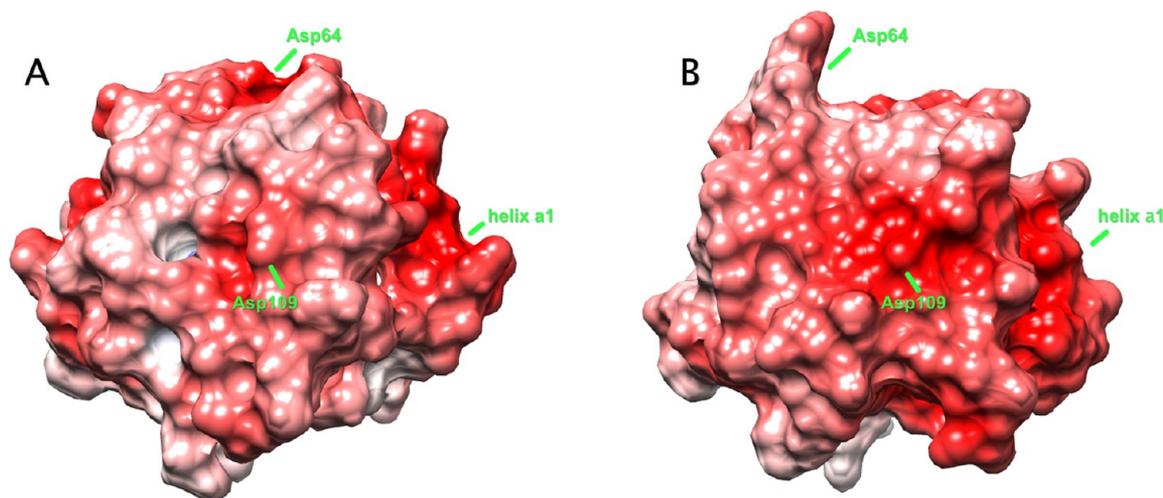
**Sidechain Changes Accompanying Activation.** An interesting question is how the phosphorylation state of the activated enzyme is communicated to the 455 face, in order to achieve the conformational transition that prepares the receiver domain to bind its cognate effector proteins. Moreover, since binding to partner molecules is likely to involve surface interactions, it is important to characterize how activation affects the surface of the receiver domain.

In typical receiver domains, active site phosphorylation propagates structural changes to the 455 face through a mechanism called Y–T coupling.<sup>2</sup> Upon phosphorylation of a conserved aspartate, a conserved Thr or Ser forms a hydrogen bond with the phosphate through its hydroxyl group. The displacement of the Thr/Ser leaves an internal cavity that is filled by the reorientation of an Tyr/Phe residue from the surface to the interior of the structure.<sup>2</sup> In Sma0114, the conserved threonine corresponds to Thr86. As shown in Figure 4A, Thr86 in strand  $\beta 4$  is in a nearly identical position before and after activation. The hydroxyl oxygen of Thr86 is 7.0 Å away from the closest fluorine atom in  $\text{BeF}_3^-$ , and therefore too far to form a bond. Another unusual feature of Sma0114 is that the conserved aromatic residue of Y–T coupling is replaced by Leu103. In contrast to Thr86, the side chain of Leu103 does undergo a reorientation upon activation. The switch, however,

is from the interior of the protein to the surface, the opposite of what is seen for the corresponding aromatic residue in typical receiver domains. These observations suggest that like the CheY2 receiver domain,<sup>12</sup> Sma0114 lacks the Y–T coupling mechanism.

On the basis of the structure of the inactive protein, we proposed that the Y–T coupling mechanism in Sma0114 may be substituted by the PFXFATGY sequence motif. This sequence motif is highly conserved in the subfamily of receiver domains that are substrates for HWE kinases. As shown in Figure 4A, the PFXFATGY sequence motif, which includes Thr86 and forms strand  $\beta 4$  of the Sma0114 structure, shows only minor structural differences between the inactive and activated protein (also see Figure 2B). Of the residues in the PFXFATGY motif, Thr86 comes closest to the  $\text{BeF}_3^-$  molecule, through its  $\text{H}\alpha$  proton, which is 4.1 Å away. Given that there are no major structural changes in the PFXFATGY motif between the inactive and activated structure, it seems unlikely that this motif plays a role in activation. Because of its conservation in the HWE kinase associated subfamily, it is possible that the PFXFATGY motif has a role in kinase binding or is part of the modified surface of Sma0114 used to bind effector molecules.

Figure 4B shows the sidechains that experience the largest changes between the inactive and activated structures. Lys105 is one of the residues with the greatest perturbation. In the



**Figure 5.** Electrostatic potential surface for the 455 face in (A) the inactive domain and (B) the activated domain. Red and blue correspond to negative and positive electrostatic potentials, respectively. The positions of Asp64, Asp109, and helix  $\alpha 1$  are indicated on the surfaces. Electrostatic potentials were calculated with the APBS server.<sup>44</sup>

inactive state, Lys105 forms a salt-bridge with Asp109 on the 455 surface of the domain. This salt bridge is broken in the activated structure, where the Lys105 side chain enters the interior of the protein to bind  $\text{BeF}_3^-$ . The repositioning of Lys105 to the interior of the domain is accompanied by the displacement of Leu103 in strand  $\beta 5$  from the interior toward the surface of the structure. The cavity left by the displacement of Leu103 is filled by Lys105 and by the aromatic ring of Phe107 from loop  $L_{\beta 5\alpha 5}$ . The movement of Phe107 leaves room for Glu111 to switch from a position on the surface exposed side of helix  $\alpha 5$  to one that bridges the gap between helices  $\alpha 5$  and  $\alpha 1$ . Part of what may contribute to bringing helix  $\alpha 1$  closer to the  $\beta$ -sheet core of the protein is an electrostatic interaction between His33 N $\epsilon$ 2 and Asp55 OD1 (not shown), which are 4.3 Å apart in the activated state but 8.2 Å in the inactive structure.

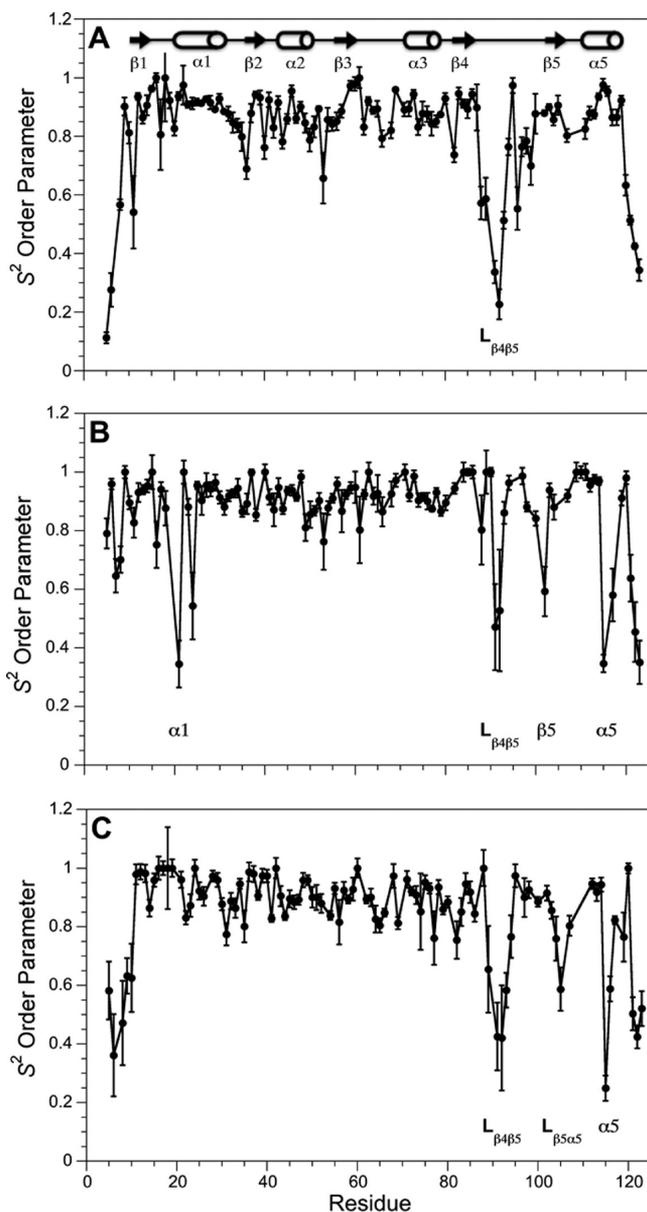
Outside of the 455 face, large changes in side chain positions are found at the active site. These include Glu17, which moves from a surface position toward the interior of the structure to act as a ligand for the  $\text{Ca}^{2+}$  ion (Figure 1C). Large changes are also seen for the entire loop  $L_{\beta 3\alpha 3}$ , spanning residues Val61–Glu66, which moves away from the active site to an “open” conformation in the activated structure (Figure 4B). The displacement of  $L_{\beta 3\alpha 3}$  in the activated structure may be due to steric occlusion from the  $\text{BeF}_3^-$  molecule,  $\text{Ca}^{2+}$  ion, and the ligating residues from the protein that move into the active site. Electrostatic repulsion between Asp64 at the center of this loop and the high density of negatively charged groups in the active site (Glu15, Asp16, Glu17, Asp60) may also contribute to the displacement of the loop.

The side chain changes accompanying activation of Sma0114 lead to changes in the surface properties of the 455 face, as shown in Figure 5. First, it is important to note that Sma0114 has a predominantly anionic character with a pI of 4.5, due to an excess of 21 acidic over 11 basic residues. Most of the surface of Sma0114 is negatively charged, except for a positively charged patch corresponding to the exposed sides of helices  $\alpha 2$  and  $\alpha 3$ , on the opposite side of the view shown in Figure 5. This unusual charge distribution is unique to Sma0114 and is not a feature shared with other receiver domains. The views in Figure 5 show the 455 face in front. The ridge to the right

below the 455 face is due to helix  $\alpha 1$ . The interface between helices  $\alpha 1$  and  $\alpha 5$  which closes the Rossmann fold structure has a high density of negatively charged residues. Activation of Sma0114 appears to increase the negative character of the 455 surface. This is particularly evident near Asp109, which is labeled in Figure 5. In the inactive state (Figure 5A), the charge on Asp109 is complemented by the salt bridge with Lys105. Upon activation, Lys105 switches to the interior of the protein to bind  $\text{BeF}_3^-$ , leaving the negative charge on Asp109 exposed (Figure 5B). Other more subtle changes such as partial burial of the basic residues Lys91 and Arg96 also contribute to the display of a more negatively charged 455 surface in the activated state (Figure 5B).

**Effects of  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$  Binding on Backbone Dynamics.** Figure 6 shows  $S^2$  order parameters, which describe the amplitudes of backbone HN motions on the ps–ns timescale, for inactive Sma0114 (Figure 6A), the Sma0114- $\text{Ca}^{2+}$  binary complex (Figure 6B), and the activated Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex (Figure 6C). The active site loops,  $L_{\beta 1\alpha 1}$  and  $L_{\beta 3\alpha 3}$ , that contribute ligands to  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$  show average  $S^2$  order parameters. The lack of ps–ns motions in these loops indicate a lack of dynamic flexibility for these regions in the inactive state (Figure 6A). As we only have data on backbone dynamics, this does not exclude increased side chain flexibility for these residues. In the inactive protein (Figure 6A), low  $S^2$  order parameters indicative of backbone flexibility occur at the chain termini and in the portion of loop  $L_{\beta 4\beta 5}$  that corresponds to helix  $\alpha 4$ . The chain termini and residues (Lys91, Gly92, Leu93) in loop  $L_{\beta 4\beta 5}$  retain low  $S^2$  terms in the all three states. Thus the region corresponding to helix  $\alpha 4$  remains unfolded in the metal-bound and activated states.

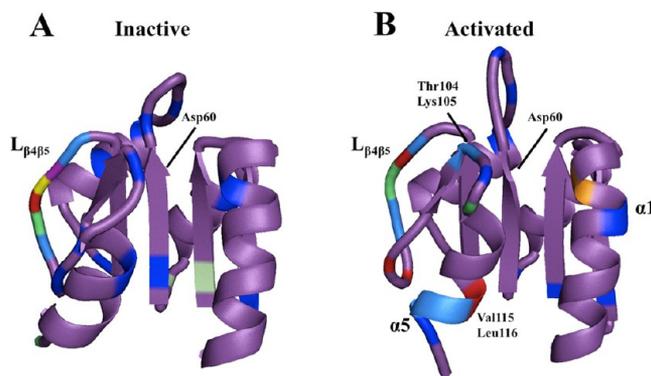
Backbone flexibility is increased at a number of sites in the Sma0114- $\text{Ca}^{2+}$  binary complex (Figure 6B) compared to the inactive protein (Figure 6A). This includes residues Met19 and Leu24 following the metal binding site. In the 455 face increased mobility is seen for residues L102 in strand  $\beta 5$  and V115–V117 at the end of helix  $\alpha 5$ . The enhanced backbone dynamics in regions of the protein remote from the binding site suggest an entropy-compensated binding mechanism,<sup>30–33</sup>



**Figure 6.** Model-free analysis of Sma0114 dynamics based on <sup>15</sup>N relaxation data. *S*<sup>2</sup> order parameters for (A) apo-Sma0114. The data for apo-Sma0114 have been published previously<sup>11</sup> and are shown here for comparison. (B) The Sma0114-Ca<sup>2+</sup> binary complex. (C) The Sma0114-Ca<sup>2+</sup>-BeF<sub>3</sub><sup>-</sup> ternary complex. The correlation times for global isotropic diffusion of the protein were 5.6 ns for A, 5.9 ns for B, and 5.7 ns for C. A secondary structure diagram with  $\alpha$ -helices represented by cylinders and  $\beta$ -sheets by arrows is shown in (A). Segments of the protein that show low *S*<sup>2</sup> order parameters are labeled in each panel.

where the ordering of the Ca<sup>2+</sup>-binding site is offset by increased mobility in other parts of the molecule.

Data on the backbone dynamics for the activated state, stabilized by the binding of Ca<sup>2+</sup> and BeF<sub>3</sub><sup>-</sup> are shown in Figure 6C. The *S*<sup>2</sup> order parameters of the inactive and activated protein are mapped on their corresponding structures in Figure 7. Mobility in helix  $\alpha$ 1 is reduced in the activated state (Figure 6C) compared to the Ca<sup>2+</sup>-bound state (Figure 6B). Like the Ca<sup>2+</sup>-bound state, the activated state shows flexibility for residues in the 455 face, although there are some subtle differences compared to the former. In the activated state,



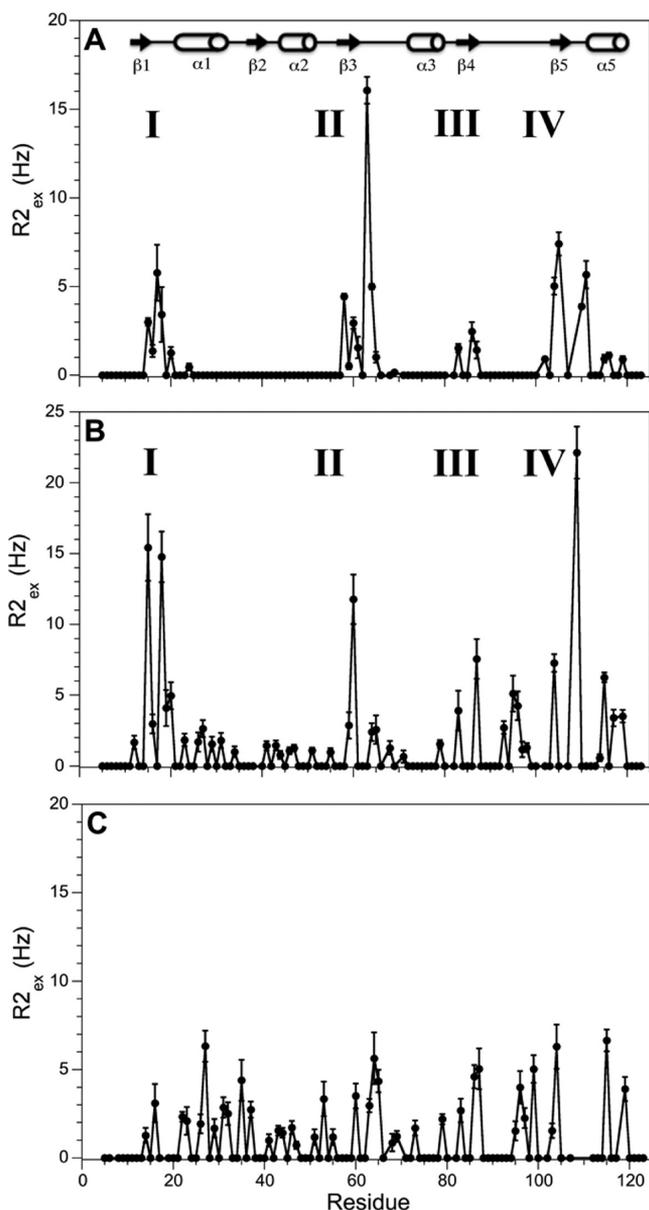
**Figure 7.** *S*<sup>2</sup> Order parameters mapped onto the NMR structures of (A) inactive and (B) activated Sma0114. The data are from Figure 6A,C. The *S*<sup>2</sup> values are represented with a ROYGBIV seven-color gradient with red being the most flexible  $0.00 \leq S^2 \leq 0.14$  and violet the least flexible  $0.84 < S^2 \leq 1.00$ . In the inactive structure, flexibility is predominantly associated with the loop  $L_{\beta 4\beta 5}$ , which replaces the canonical receiver domain fold  $\alpha$ 4 helix in Sma0114 (A). In the activated domain, the loop  $L_{\beta 4\beta 5}$  remains flexible and additional flexibility is seen for residues in helix  $\alpha$ 1,  $\alpha$ 5, and loop  $L_{\beta 5\alpha 5}$  with the latter two belonging to the 455-face of the domain. Residues experiencing the greatest change in flexibility, Thr104, Lys105, Val115 and Leu116, are labeled on the activated structure. The phosphorylation site, Asp60, is also labeled.

residues Thr104-Phe107 in loop  $L_{\beta 5\alpha 5}$  show raised mobility, whereas in the Ca<sup>2+</sup>-bound state residue Leu102 in strand  $\beta$ 5 is mobile (Figure 6B,C). Interestingly, the residue with the highest backbone mobility in loop  $L_{\beta 5\alpha 5}$  of the activated state is Lys105, whose side chain binds BeF<sub>3</sub><sup>-</sup>. Like the Ca<sup>2+</sup>-bound protein, the activated state exhibits low *S*<sup>2</sup> order parameters for residues Val115-Val117 at the C-terminal end of helix  $\alpha$ 5.

Figure 8 shows R<sub>2ex</sub> line-broadening contributions for the three states obtained from Model Free calculations.<sup>25–27</sup> In both the apo protein (Figure 8A) and in the Ca<sup>2+</sup>-bound state (Figure 8B), there are four regions that have large R<sub>2ex</sub> terms indicative of conformational exchange on the  $\mu$ s–ms time scale. These are (I) the metal binding loop ( $L_{\beta 1\alpha 1}$ ), (II)  $L_{\beta 3\alpha 3}$  following the site of phosphorylation Asp60, (III) the C-terminal region of strand  $\beta$ 4, and (IV) the  $\beta$ 5- $L_{\beta 5\alpha 5}$  segment that holds Lys105, which stabilizes the acyl-phosphate moiety. R<sub>2ex</sub> contributions for these regions have been described in other receiver domains and have been attributed to a dynamic equilibrium between the inactive and active conformations.<sup>5,34</sup> Consistent with previous work on other receiver domains,<sup>5,34</sup> the R<sub>2ex</sub> terms for regions I–IV in Sma0114 decrease as the equilibrium is shifted in favor of the activated state, by binding Ca<sup>2+</sup> and BeF<sub>3</sub><sup>-</sup> (Figure 8C). The large number of R<sub>2ex</sub> terms between 0 and 5 Hz for the activated state is an artifact of lower precision <sup>15</sup>N relaxation data. The relatively high salt concentration of 30 mM NaF, needed to generate BeF<sub>3</sub><sup>-</sup> in the sample, reduces the sensitivity of the relaxation data. A second factor that reduces the precision of the <sup>15</sup>N-relaxation data for the activated state is that the ternary complex is more subject to precipitation during the long periods (~1 week) needed to perform the <sup>15</sup>N-relaxation experiments.

## DISCUSSION

Phosphorylation causes a conformational switch that enables a receiver domain to bind downstream target molecules.<sup>4</sup> The duration of the output response, which can range from seconds



**Figure 8.**  $R2_{ex}$  contributions derived from Model Free analysis of  $^{15}\text{N}$ -relaxation data for (A) apo-Sma0114. The data for apo-Sma0114 have been published previously<sup>11</sup> and are shown here for comparison. (B) The Sma0114- $\text{Ca}^{2+}$  binary complex. (C) The Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex. The secondary structure is shown in (A). The inactive apo-domain (A) shows four regions with raised  $R2_{ex}$  values, indicative of  $\mu\text{s}$ - $\text{ms}$  conformational exchange: (I) the metal-binding site, (II) the phosphorylation site, (III) the PFxFATGY motif that forms strand  $\beta 4$ , and (IV) the junction between strand  $\beta 5$  and helix  $\alpha 5$ . The  $R2_{ex}$  contributions persist when  $\text{Ca}^{2+}$  is bound (B) but are suppressed in the activated state of the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex.

to hours, is regulated by the lifetime of the acyl phosphate linkage that induces the conformational transition to the activated state. In the current study, we stabilized the activated state of Sma0114 by binding  $\text{Ca}^{2+}$  and the phosphate analogue  $\text{BeF}_3^-$ , as has been done in other studies of receiver domains<sup>12,14</sup> due to the limited lifetime of the acyl phosphate. We previously reported the NMR structure of the inactive form of Sma0114, which plays a role in metabolic regulation and catabolite repression in the bacterium *Sinorhizobium meliloti*.<sup>11</sup>

In the present work, we describe the NMR structure and dynamics of the activated form of Sma0114 to provide a more complete picture of the activation mechanism of this response regulator that is prototypical for the family of receiver domains activated by HWE histidine kinases.

A large difference between the structures of inactive and activated Sma0114 occurs in the loop  $L_{\beta 3\alpha 3}$ . In the inactive state, this loop folds over the active site, partially occluding it (Figure 2A). This structural feature may confer specificity, for example, guarding the active site against phosphorylation by noncognate kinases. Our active state structure shows that once phosphorylation occurs at Asp60,  $L_{\beta 3\alpha 3}$  switches to an open conformation that leaves the active site more exposed compared to the inactive state (Figure 2A). This open  $L_{\beta 3\alpha 3}$  loop conformation appears to be conserved in the activated states of other receiver domains (Figure 3). A possible advantage of having the  $L_{\beta 3\alpha 3}$  loop displaced from the active site is that it could facilitate the release of the phosphate leaving group, after hydrolysis by the autophosphatase activity of the receiver domain. As such, loop  $L_{\beta 3\alpha 3}$  could participate in determining the lifetime of the activated state before the enzyme switches back to its inactive state. Other factors that can modulate the lifetime of the activated state include the types and concentrations of intracellular metal ions, since these affect the enzymatic activity of the receiver domain.<sup>6,11</sup> Receiver domains bind a number of divalent metal cations including  $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Ca}^{2+}$ , and  $\text{Cu}^{2+}$ .<sup>6,35</sup> On the basis of NMR titration studies, we found that the  $K_d$  values for metal binding to Sma0114 were  $\sim 75$  mM for  $\text{Mg}^{2+}$  and  $\sim 1$  mM for  $\text{Ca}^{2+}$ .<sup>15</sup> The high ionic strength conditions required for  $\text{Mg}^{2+}$  studies, nonspecific binding at the high 150 mM  $\text{Mg}^{2+}$  concentrations needed to saturate the active site, together with an enhanced propensity for protein aggregation under these conditions, led us to choose  $\text{Ca}^{2+}$  as the divalent metal ion to stabilize the activated state, rather than the more typical  $\text{Mg}^{2+}$ . Under saturating conditions, the active site chemical shifts for the  $\text{Mg}^{2+}$ - and  $\text{Ca}^{2+}$ -bound forms of Sma0114 were very similar, suggesting similar structures for the metal bound states, consistent with similar crystallographic structures for receiver domains such as Spo0F bound by different divalent metals.<sup>35,36</sup> The  $K_d$  values for both  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  binding by Sma0114 are about 100-fold larger than the physiological concentrations of the metals. A very weak affinity for divalent metals is a characteristic of other receiver domains, such as Spo0F and DivK.<sup>35</sup> It has been proposed that this weak metal affinity has a regulatory role. A weak  $K_d$  compared to the physiological metal concentrations would lower the autophosphatase activity of the receiver domain thereby extending the lifetime of the activated state.<sup>6,36</sup>

A key question in the structural biology of receiver domains is how phosphorylation of the conserved aspartate induces the conformational switch needed to bind downstream effector molecules. Typically the initial conformational switch occurs through Y-T coupling when a threonine at the C-terminal end of strand  $\beta 4$  repositions to hydrogen bond to the phosphate. The movement of the threonine allows the reorientation of an aromatic residue in strand  $\beta 5$  from the surface of the 455 face to the interior of the protein, thus propagating conformational changes from the active site to the 455 face.<sup>2</sup> In Sma0114, Thr86, corresponding to the threonine in Y-T coupling, is too far from the active site to hydrogen bond to the phosphate and shows little structural difference between the inactive and activated states (Figure 4A). The conserved aromatic residue is

replaced by Leu103, which switches from the inside to the outside of the Sma0114 structure on activation, the opposite of what is seen for the aromatic residue in classical Y–T coupling. Receiver domains that are substrates for HWE kinases have a highly conserved PFxFATGY sequence motif. Because this motif encompasses Thr86, corresponding to the threonine of Y–T coupling, we proposed that this motif could be used as an alternative for communication between the active site and the 455 face in Sma0114.<sup>11</sup> The NMR structure of the activated state, however, shows that residues of the PFxFATGY motif in strand  $\beta 4$  retain similar conformations to the inactive state and are too far away from the active site to substitute for the Y–T coupling mechanism (Figure 4A). We conclude that Y–T coupling is eliminated in Sma0114 and that the conserved PFxFATGY motif is not an alternative conduit of communication between the active site and the 455 face used to bind downstream effector molecules.

The only interaction between the 455 face and the active site that to our knowledge is conserved in all receiver domains is that between Lys105 in loop  $L_{\beta 5\alpha 5}$  and the phosphate or  $\text{BeF}_3^-$  ion. In the activated state of Sma0114, the repositioning of Lys105 to bind  $\text{BeF}_3^-$  is accompanied by a number of changes that affect the 455 face. These include a displacement of Leu103 in strand  $\beta 5$  from the interior to the surface of the protein, and movements of Phe107 and Glu111 toward the interior of the protein that bring helix  $\alpha 5$  closer to the  $\beta$ -sheet core and to helix  $\alpha 1$  in the activated structure. Although the role of Lys105 has traditionally been viewed as stabilizing the acyl phosphate,<sup>3</sup> it also seems the most likely candidate to effectuate the communication between the active site and 455 face of Sma0114.

Because different receiver domains recognize a variety of downstream targets, the structural changes accompanying activation are variable. In most receiver domains, the largest changes induced by activation occur on the 455 face (Figure 3); however, this is not always the case. Activation of the bacterial sporulation factor Spo0F leaves the 455 face relatively unchanged. The largest structural changes occur in helix  $\alpha 1$  (Figure 3), and these have been shown to play important roles in protein–protein interactions with the downstream targets of Spo0F.<sup>17</sup> Activation of Sma0114 leads to structural changes in strand  $\beta 5$  and helix  $\alpha 5$  of the 455 face (Figure 2), with the important difference that helix  $\alpha 4$  of the typical receiver domain fold is replaced by a flexible loop. Additional changes extend to helix  $\alpha 1$ , which is also perturbed in the activated state of the three receiver domains NtrC, CheY, and Spo0F (Figure 3). We cannot tell if the rearrangement of helix  $\alpha 1$  in Sma0114 is due to  $\text{Ca}^{2+}$  binding or requires both  $\text{Ca}^{2+}$  and  $\text{BeF}_3^-$ . Therefore, helix  $\alpha 1$  may have a role in priming the enzyme for phosphorylation or participate in the output response by binding effectors in the fully activated state. Changes in secondary structure elements accompanying activation must lead to changes in the surfaces of receiver domains that enable them to bind their downstream target molecules. In the case of Sma0114, the side chain rearrangements that accompany activation give the 455 face a more negatively charged character (Figure 5) that could be used to select positively charged target molecules.

Receiver domain dynamics on slow  $\mu\text{s}$ – $\text{ms}$  timescales have been extensively investigated using relaxation dispersion methods.<sup>5,34,37</sup> These studies showed that the inactive and activated conformations exist in a dynamic equilibrium, that is shifted in favor of the activated state when the active site binds

ligands.<sup>5</sup> Our results for Sma0114 are consistent with this two-state allosteric mechanism of activation. The apo (Figure 8A) and  $\text{Ca}^{2+}$ -bound enzyme (Figure 8B) exhibit  $R_{2,\text{ex}}$  line-broadening contributions for NMR signals from residues in the active site and 455 face. The  $R_{2,\text{ex}}$  contributions are suppressed in the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex, where the population of the activated state dominates (Figure 8C).

By contrast, there have been few studies of receiver domain dynamics on the fast ps–ns timescales accessible by  $^{15}\text{N}$  relaxation measurements. Spo0F was characterized only in the apo-state where dynamics on the ps–ns time scale were shown to be constant except at the chain termini.<sup>38</sup> For NtrC, ps–ns dynamics were characterized for both the inactive and activated forms of the enzyme and were shown to be constant except at the chain termini.<sup>5</sup> In the case of Sma0114, we characterized ps–ns dynamics for three forms of the protein: the apo-state, the  $\text{Ca}^{2+}$ -bound state, and the activated state (Figure 6). Both the Sma0114- $\text{Ca}^{2+}$  binary complex (Figure 6B) and the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex (Figure 6C) show an increase in the number of residues exhibiting flexibility on the ps–ns time scale compared to the apoprotein (Figure 6A). These results suggest an entropy-compensated binding mechanism for the  $\text{BeF}_3^-$  phosphate analogue, where the decrease in entropy due to ligand binding in the active site is compensated by increased dynamics in other parts of the structure. Entropy-compensated binding was first hypothesized in 1963 for insulin dimerization as a way to alleviate the loss of monomer rotational and translational degrees of freedom by an increase in vibrational dynamics.<sup>30</sup> The mechanism has since been supported by NMR relaxation studies showing increased dynamics accompanying ligand binding in a number of proteins,<sup>31,32,39</sup> including the enzymes MutT,<sup>33</sup> stromelysin,<sup>40</sup> and topoisomerase I.<sup>41</sup>

In the case of Sma0114, entropy-compensated binding is likely to be exploited for a functional role, as most of the residues that show increased dynamics in the activated state are part of the 455 face of the protein. These include residues Thr104-Phe107 in strand  $\beta 5$  and loop  $L_{\beta 5\alpha 5}$ , and Val115-Val117 in helix  $\alpha 5$ . The mobile loop (Gly89-Leu93) that replaces helix  $\alpha 4$ , together with the above segments, gives the 455 face in Sma0114 a flexible character (Figure 7B). Flexibility could be used in an induced-fit binding mechanism that would allow the 455 face to adapt to multiple downstream effector proteins. In this regard it is interesting to note that genetic studies of *sma0114* mutants suggest that the Sma0114 receiver domain indeed recognizes multiple target molecules.<sup>10</sup>

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

Two figures showing  $^{15}\text{N}$   $R_1$ ,  $R_2$ , and NOE relaxation data for the Sma0114- $\text{Ca}^{2+}$  binary complex and the Sma0114- $\text{Ca}^{2+}$ - $\text{BeF}_3^-$  ternary complex. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Address: Department of Molecular and Cell Biology, University of Connecticut, 91 N. Eagleville Rd., Storrs, CT 06269-3125, USA. Phone: (860) 486-4414. Fax: (860) 486-4331. E-mail: [andrei@uconn.edu](mailto:andrei@uconn.edu).

## Funding

This work was supported by a National Science Foundation Graduate Research Fellowship (NSF-GRFP) to S.R.S, a grant from the UConn Research Foundation (UCRF) to A.T.A. and D.J.G. and by DOE grants (DE-FG02-01ER15175 and DE-FG02-06ER15805) to D.J.G.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Prof. Victoria Robinson for useful discussions.

## ABBREVIATIONS

HWE, histidine-tryptophan-glutamate; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy;  $R2_{\text{ex}}$ , chemical exchange contribution to transverse ( $R2$ ) relaxation; RMS, root-mean-square; RMSD, root-mean-square deviation

## REFERENCES

- (1) Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction. *Annu. Rev. Biochem.* 69, 183–215.
- (2) Robinson, V. L., Buckler, D. R., and Stock, A. M. (2000) A tale of two components: a novel kinase and a regulatory switch. *Nat. Struct. Biol.* 7, 626–633.
- (3) Bourret, R. B. (2010) Receiver domain structure and function in response regulator proteins. *Curr. Opin. Microbiol.* 13, 142–149.
- (4) Jenal, U., and Galperin, M. Y. (2009) Single domain response regulators: molecular switches with emerging roles in cell organization and dynamics. *Curr. Opin. Microbiol.* 12, 152–160.
- (5) Volkman, B. F., Lipson, D., Wemmer, D. E., and Kern, D. (2001) Two-state allosteric behavior in a single-domain signaling protein. *Science* 291, 2429–2433.
- (6) Kojetin, D. J., Thompson, R. J., Benson, L. M., Naylor, S., Waterman, J., Davies, K. G., Opperman, C. H., Stephenson, K., Hoch, J. A., and Cavanagh, J. (2005) Structural analysis of divalent metals binding to the *Bacillus subtilis* response regulator Spo0F: the possibility for in vitro metalloregulation in the initiation of sporulation. *Biometals* 18, 449–466.
- (7) Wright, G. D., Holman, T. R., and Walsh, C. T. (1993) Purification and characterization of VanR and the cytosolic domain of VanS: a two-component regulatory system required for vancomycin resistance in *Enterococcus faecium* BM4147. *Biochemistry* 32, 5057–5063.
- (8) Hess, J. F., Oosawa, K., Kaplan, N., and Simon, M. I. (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* 53, 79–87.
- (9) Karniol, B., and Vierstra, R. D. (2004) The HWE histidine kinases, a new family of bacterial two-component sensor kinases with potentially diverse roles in environmental signaling. *J. Bacteriol.* 186, 445–453.
- (10) Garcia, P. P., Bringham, R. M., Arango Pinedo, C., and Gage, D. J. (2009) Characterization of a two-component regulatory system that regulates succinate-mediated catabolite repression in *Sinorhizobium meliloti*. *J. Bacteriol.* 192, 5725–5735.
- (11) Sheftic, S. R., Garcia, P. P., White, E., Robinson, V. L., Gage, D. J., and Alexandrescu, A. T. (2012) Nuclear magnetic resonance structure and dynamics of the response regulator Sma0114 from *Sinorhizobium meliloti*. *Biochemistry* 51, 6932–6941.
- (12) Riepl, H., Scharf, B., Schmitt, R., Kalbitzer, H. R., and Maurer, T. (2004) Solution structures of the inactive and BeF<sub>3</sub>-activated response regulator CheY2. *J. Mol. Biol.* 338, 287–297.
- (13) Yan, D., Cho, H. S., Hastings, C. A., Igo, M. M., Lee, S. Y., Pelton, J. G., Stewart, V., Wemmer, D. E., and Kustu, S. (1999) Beryllifluoride mimics phosphorylation of NtrC and other bacterial response regulators. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14789–14794.
- (14) Wemmer, D. E., and Kern, D. (2005) Beryllifluoride binding mimics phosphorylation of aspartate in response regulators. *J. Bacteriol.* 187, 8229–8230.
- (15) Sheftic, S. R., Garcia, P. P., Robinson, V. L., Gage, D. J., and Alexandrescu, A. T. (2010) NMR assignments for the *Sinorhizobium meliloti* response regulator Sma0114. *Biomol. NMR Assign.* 5, 55–58.
- (16) Sheftic, S. R., Garcia, P. P., White, E., Robinson, V. L., Gage, D. J., and Alexandrescu, A. T. (2012) Nuclear magnetic resonance structure and dynamics of the response regulator Sma0114 from *Sinorhizobium meliloti*. *Biochemistry* 51, 6932–6941.
- (17) Gardino, A. K., Volkman, B. F., Cho, H. S., Lee, S. Y., Wemmer, D. E., and Kern, D. (2003) The NMR solution structure of BeF<sub>3</sub>(–)-activated Spo0F reveals the conformational switch in a phosphorelay system. *J. Mol. Biol.* 331, 245–254.
- (18) Hastings, C. A., Lee, S. Y., Cho, H. S., Yan, D., Kustu, S., and Wemmer, D. E. (2003) High-resolution solution structure of the beryllifluoride-activated NtrC receiver domain. *Biochemistry* 42, 9081–9090.
- (19) Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, Rance, M., and Skelton, N. J. (2006) *Protein NMR Spectroscopy Principles and Practice (Second Edition)*, Elsevier Inc., Amsterdam.
- (20) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59, 687–696.
- (21) Brunger, A. T. (1996) Xplor Version 3.8, A System for Crystallography and NMR, Yale University Press, New Haven.
- (22) Cates, M. S., Teodoro, M. L., and Phillips, G. N., Jr. (2002) Molecular mechanisms of calcium and magnesium binding to parvalbumin. *Biophys. J.* 82, 1133–1146.
- (23) Stites, W. E., Gittis, A. G., Lattman, E. E., and Shortle, D. (1991) In a staphylococcal nuclease mutant the side-chain of a lysine replacing valine 66 is fully buried in the hydrophobic core. *J. Mol. Biol.* 221, 7–14.
- (24) Nicholson, L. K., Kay, L. E., and Torchia, D. A. (1994) Protein Dynamics as studied by solution NMR techniques., In *NMR Spectroscopy and its Application to Biomedical Research* (Sarkar, S., Ed.), Elsevier Science Publishers.
- (25) Alexandrescu, A. T., and Shortle, D. (1994) Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease. *J. Mol. Biol.* 242, 527–546.
- (26) Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Validity. *J. Am. Chem. Soc.* 104, 4546–4559.
- (27) Dosset, P., Hus, J. C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* 16, 23–28.
- (28) Madhusudan, M., Zapf, J., Hoch, J. A., Whiteley, J. M., Xuong, N. H., and Varughese, K. I. (1997) A response regulatory protein with the site of phosphorylation blocked by an arginine interaction: crystal structure of Spo0F from *Bacillus subtilis*. *Biochemistry* 36, 12739–12745.
- (29) Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M., and Laub, M. T. (2008) Rewiring the specificity of two-component signal transduction systems. *Cell* 133, 1043–1054.
- (30) Steinberg, I. Z., and Scheraga, H. A. (1963) Entropy changes accompanying association reactions of proteins. *J. Biol. Chem.* 238, 172–181.
- (31) Zidek, L., Novotny, M. V., and Stone, M. J. (1999) Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nat. Struct. Biol.* 6, 1118–1121.
- (32) Arumugam, S., Gao, G., Patton, B. L., Semenchenko, V., Brew, K., and Van Doren, S. R. (2003) Increased backbone mobility in beta-barrel enhances entropy gain driving binding of N-TIMP-1 to MMP-3. *J. Mol. Biol.* 327, 719–734.
- (33) Stivers, J. T., Abeygunawardana, C., and Mildvan, A. S. (1996) 15N NMR relaxation studies of free and inhibitor-bound 4-

oxalocrotonate tautomerase: backbone dynamics and entropy changes of an enzyme upon inhibitor binding. *Biochemistry* 35, 16036–16047.

(34) Gardino, A. K., and Kern, D. (2007) Functional dynamics of response regulators using NMR relaxation techniques. *Methods Enzymol.* 423, 149–165.

(35) Mukhopadhyay, D., Sen, U., Zapf, J., and Varughese, K. I. (2004) Metals in the sporulation phosphorelay: manganese binding by the response regulator Spo0F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 60, 638–645.

(36) Madhusudan, Z. J., Whiteley, J. M., Hoch, J. A., Xuong, N. H., and Varughese, K. I. (1996) Crystal structure of a phosphatase-resistant mutant of sporulation response regulator Spo0F from *Bacillus subtilis*. *Structure* 4, 679–690.

(37) Feher, V. A., Zapf, J. W., Hoch, J. A., Whiteley, J. M., McIntosh, L. P., Rance, M., Skelton, N. J., Dahlquist, F. W., and Cavanagh, J. (1997) High-resolution NMR structure and backbone dynamics of the *Bacillus subtilis* response regulator, Spo0F: implications for phosphorylation and molecular recognition. *Biochemistry* 36, 10015–10025.

(38) Feher, V. A., and Cavanagh, J. (1999) Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. *Nature* 400, 289–293.

(39) Stone, M. J. (2001) NMR relaxation studies of the role of conformational entropy in protein stability and ligand binding. *Acc. Chem. Res.* 34, 379–388.

(40) Yuan, P., Marshall, V. P., Petzold, G. L., Poorman, R. A., and Stockman, B. J. (1999) Dynamics of stromelysin/inhibitor interactions studied by <sup>15</sup>N NMR relaxation measurements: comparison of ligand binding to the S1-S3 and S'1-S'3 subsites. *J. Biomol. NMR* 15, 55–64.

(41) Yu, L., Zhu, C. X., Tse-Dinh, Y. C., and Fesik, S. W. (1996) Backbone dynamics of the C-terminal domain of *Escherichia coli* topoisomerase I in the absence and presence of single-stranded DNA. *Biochemistry* 35, 9661–9666.

(42) Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4, 187–217.

(43) Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.

(44) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10037–10041.