HHS Public Access

Author manuscript

Biochemistry. Author manuscript; available in PMC 2016 February 17.

Published in final edited form as:

Biochemistry. 2015 February 17; 54(6): 1353–1363. doi:10.1021/bi501143b.

Ligand-induced folding of a two component signaling receiver domain

Victor J. Ocasio[1], **Fernando Corrêa**[1,2], and **Kevin H. Gardner**[1,2,3]

Kevin H. Gardner: Kevin.Gardner@asrc.cuny.edu

[1]Departments of Biophysics and Biochemistry, UT Southwestern Medical Center, Dallas, TX 75390-8816

[2]Structural Biology Initiative, CUNY Advanced Science Research Center, New York, NY 10031

[3]Department of Chemistry, City College of New York, New York, NY 10031

Abstract

To survive and adapt to environmental changes, bacteria commonly use two component signaling systems. Minimally, these pathways use histidine kinases (HKs) to detect environmental signals, harnessing these to control phosphorylation levels of receiver (REC) domains of downstream response regulators that convert this signal into physiological responses. Studies of several prototypical REC domains suggest that phosphorylation shifts these proteins between inactive and active structures that are globally similar and well-folded. However, it is unclear how globally these findings hold within REC domains in general, particularly when considered within full-length proteins. Here we present EL_LovR, a full-length REC-only protein that is phosphorylated in response to blue light in the marine α-proteobacterium *Erythrobacter litoralis* HTCC2594. Notably, EL_LovR is similar to comparable REC-only proteins used in bacterial general stress responses, where genetic evidence suggests that their potent phosphatase activity is important to shut off such systems. Size exclusion chromatography, light scattering and solution NMR experiments show that EL_LovR is monomeric and unfolded in solution under conditions routinely used for other REC structure determinations. Addition of Mg^{2+} and phosphorylation induce progressively greater degrees of tertiary structure stabilization, with the solution structure of the fully-activated EL_LovR adopting the canonical receiver domain fold. Parallel functional assays show that EL_LovR has a fast dephosphorylation rate, consistent with its proposed function as a phosphate sink that depletes the HK phosphoryl group, promoting the phosphatase activity of this enzyme. Our findings demonstrate that EL_LovR undergoes substantial ligand-dependent conformational changes that have not been reported for other RRs, expanding the scope of conformational changes and regulation used by REC domains, critical components of bacterial signaling systems.

Correspondence to: Kevin H. Gardner, Kevin.Gardner@asrc.cuny.edu.

SUPPORTING INFORMATION

Includes five Supporting Figures presenting ^{15}N-edited NOESY spectra of apo and Mg^{2+}/BeF_3^- bound EL_LovR (Figure S1), ^{2}H exchange and ^{15}N NMR relaxation measurements of Mg^{2+}/BeF_3^- bound EL_LovR (Figure S2), $^{15}N/^{1}H$ HSQC spectra of EL_LovR D56A with various ligands bound (Figure S3), SEC-MALS data of EL_LovR D13A and D56A mutants (Figure S4) and sequence alignment between EL_LovR and other REC domains (Figure S5). This material is available free of charge via the Internet at http://pubs.acs.org.

Two-component signal transduction (TCS) systems are the most prevalent strategy used by bacteria to sense and adapt to changes in their environment[1,2]. Minimally, TCS are comprised of a sensor histidine kinase (HK) and a response regulator (RR)[3]. HKs typically contain three types of domains: an environmental sensor, a dimerization and histidine phosphotransfer domain (DHp) and a catalytic domain (CA). Their combined operation allows an HK protein to sense environmental cues via the sensor domain and translate this signal into changes in phosphorylation level on a critical His residue in the DHp domain. With the help of a Mg^{2+} ion, the phosphoryl group is transferred from the phospho-His residue to an aspartate in the receiver domain (REC) of the downstream RR, controlling its function.

While all REC domains share a conserved $(\beta\alpha)_5$ fold (Figure 1) and phosphoacceptor region that includes the critical phosphorylated aspartate and several nearby acidic residues required for binding Mg^{2+} ion[4], these domains are found in a wide variety of protein architectures. Some REC-containing proteins contain different types of effectors (e.g. DNA binding domains), which are directly controlled by phosphorylation, while others contain solely isolated REC domains. This latter group, collectively referred to as single domain response regulators (SDRRs), are fairly prevalent, composing the second largest class of RR proteins (~14%)[5,6]. While these proteins lack an effector domain of their own, they can use the α4-β5-α5 surface at their C-termini to regulate functions of many other diverse proteins. This often occurs by activation-controlled protein/protein interactions; for example, when the CheY SDRR chemotaxis protein is phosphorylated, it interacts with a member of the switch of flagellar motor FliM[7], changing the direction of flagellar rotation. Additionally, it has been reported that CheY can also function as a phosphate sink[5]. Another SDRR, DivK, plays an essential role in *Caulobacter crescentus* cell division by temporally regulating proteolysis of CtrA[8], a RR that regulates the expression of many genes involved in cell cycle[9]. The sole common theme among these functionally distinct proteins is the REC domain fold.

The widespread use of REC domains in bacterial signaling has led to intense interest in understanding how phosphorylation activates these switches and thereby controls their function. To address these questions, a number of REC domain structures have been solved in their active[10–14] and inactive states[15–18]and used to generate models of REC signaling. One such model entails the use of phosphorylation to shift a preexisting structural equilibrium, as perhaps best validated by data collected on the REC domain of NtrC[19]. When unphosphorylated, this REC domain rapidly interconverts between well-structured inactive and active-like conformations, with the equilibrium significantly favoring the lower energy, inactive conformation. Upon phosphorylation, the equilibrium shifts to fully populate the active state[19]. While elegant, the generality of this signaling model remains somewhat unclear, as some REC domains clearly undergo non-two state segmental sampling of active-like conformations (CheY[20]) or partial unfolding of different secondary structure elements (Sma0114[14, 18]). These examples underscore the need to more completely characterize different REC domain signaling mechanisms in light of the apparent diversity that has been revealed to date.

To this end, we integrated biochemical and biophysical approaches to characterize activation-induced changes in EL_LovR, an SDRR involved in a light regulated TCS[21] found in the marine α-proteobacterium *Erythrobacter litoralis*. EL_LovR, one of the 23 predicted RR encoded in the *E. litoralis* genome[21], is one of two RRs phosphorylated by three light-sensitive HKs in *E. litoralis*, EL346/EL362/EL368[21, 22]. Coupling these data with bioinformatics and genetics analyses showing the importance of a close homolog (LovR) in stress responses in *Caulobacter crescentus*[23, 24], we hypothesized that EL_LovR has a similar role and sought to investigate how it might be affected by phosphorylation.

Our biophysical analyses revealed that EL_LovR fundamentally requires two ligands, Mg^{2+} ion and phosphorylation, to stably fold into the standard REC domain structure. Contrary to expectations that EL_LovR would adopt a canonical REC structure on its own, size exclusion chromatography, limited proteolysis and solution NMR spectroscopy all showed that ligand-free EL_LovR is substantially disordered with limited secondary and tertiary structure. Binding Mg^{2+} ion led to partial stabilization of the REC domain fold, which was increased by enzymatic or small molecule phosphorylation or binding to the phosphomimetic beryllium fluoride (BeF_3^-). Using solution NMR spectroscopy, we further confirmed that activated EL_LovR bound to both Mg^{2+} and BeF_3^- adopts the conserved REC domain fold as originally anticipated. Additionally, phosphatase assays indicate that the active state of EL_LovR is short lived with a half-life of approximately 2 min, consistent with its likely role as a phosphate sink like its *C. crescentus* homolog LovR[23]. These data establish that EL_LovR undergoes a substantial ligand-dependent folding reaction, perhaps ensuring that its phosphatase activity is restricted to specific circumstances.

Experimental procedures

Cloning, expression, and purification of LOV-HK and response regulator proteins

DNA encoding sequences of the LOV-HK protein EL368, (NCBI Gene locus tag ELI_02980) were amplified from *Erythrobacter litoralis* HTCC2594 genomic DNA[25], and cloned into the pHis-Gβ1-parallel expression vector[26]. Response regulators EL_LovR and EL_PhyR (NCBI gene locus tags ELI_07655 and ELI_10215) were amplified from *Erythrobacter litoralis* HTCC2594 genomic DNA and cloned into the pHis-parallel expression vector[26]. EL_LovR mutants (E12A, D13A, E14A, D56A, P95A, N54K) were generated by site directed mutagenesis using primers encoding the desired mutation. Protein production in complex media and purification were performed as described[21], with protein concentrations measured using Bradford assays (Pierce)[27]. U-^{15}N labeled and ^{15}N,^{13}C labeled proteins were obtained by transforming protein expression plasmids into *Escherichia coli* BL21(DE3) cells grown in M9 minimal media containing 1 g/L of $^{15}NH_4Cl$ for U-^{15}N samples, supplemented with 3 g/L [$^{13}C_6$] glucose for U-^{15}N/^{13}C labeled samples. Cultures were shaken at 37 °C to an A_{600} of 0.6–0.8, before gene expression was induced with 0.5 mM isopropyl β-D-thiogalactopiranoside at 18 °C. Protein purifications were conducted as previously reported[21].

Size exclusion chromatography and multiple laser light scattering

The oligomerization state of EL_LovR was determined by size exclusion chromatography coupled to inline multiangle light scattering (SEC-MALS). 500 μL of 20 μM EL_LovR was injected onto a Superdex 200 10/300 analytical gel filtration column (GE Biosciences) pre-equilibrated with 50 mM Tris (pH 8.0), 100 mM NaCl, 5 mM DTT buffer. In conditions that required magnesium, 10 mM $MgCl_2$ was added to sample and column buffers. 10 mM carbamoyl phosphate and 10 mM $MgCl_2$ was incubated with EL_LovR for 10 min to generate the phosphorylated state. Elution volumes were calibrated to apparent molecular weights using the following standards: thyroglobulin (670 kDa), gamma-globulin (158 kDa), ovalbumin (44 kDa), myoglobulin (17 kDa) and vitamin B12 (1.4 kDa). Post-chromatography, eluting protein was detected using inline miniDAWN TREOS light scattering and Optilab rEX refractive index detectors (Wyatt Technology). All sample collection was done at 4 °C. Data analyses and molecular weight calculations were carried out using the ASTRA V software (Wyatt Technology).

Limited proteolysis and mass spectrometry analyses

Limited proteolysis was carried out by mixing EL_LovR (390 μM = 5 mg/mL) with trypsin (w/w 1 trypsin: 90 EL_LovR) in 50 mM Tris (pH 7.5), 100 mM NaCl, 5 mM DTT. For experiments requiring Mg^{2+} and BeF_3^-, 10 mM of $MgCl_2$ and 5 mM/15 mM of $BeCl_2$/NaF were added, respectively. After mixing EL_LovR and trypsin, 10 μL aliquots were taken at 1, 3, 6, 10 and 15 min timepoints and stopped with 4× SDS–PAGE sample buffer (50 mM Tris (pH 6.8), 200 mM NaCl, 40 mM EDTA, 0.2% bromophenol blue, 10% (v/v) β-mercaptoethanol, 4% (w/w) SDS, and 20% (v/v) glycerol) for gel electrophoresis analysis. For mass spectrometry analysis, EL_LovR was digested for 3 min and the reaction was stopped with 4 mM phenylmethylsulfonyl fluoride (PMSF). Molecular masses of resulting fragments were obtained by LC-MS with Agilent 6540 Q-TOF instrument (UT Southwestern Proteomics Core Facility).

Phosphorylation and phosphatase assays

EL_LovR and EL_PhyR (20 μM) were incubated with EL368 (10 μM) in 50 mM Tris (pH 7.5), 100 mM NaCl, 5 mM DTT at room temperature. Proteins were incubated for 10 min at room temperature with 0.5 mM ATP and 25 μCi [γ-^{32}P] ATP (6000 Ci/mmol, PerkinElmer), with 10 μL aliquots taken at 0.5, 1, 5 and 10 min intervals. Samples were treated as previously described[21]. Phosphatase assays were done by incubating 5 μM EL368 and 20 μM EL_LovR in 10 mM Tris (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT. Proteins were incubated with 500 μM ATP and 70 μCi [γ-^{32}P] ATP (6000 Ci/mmol, PerkinElmer) for 10 min at room temperature. After this 10 min incubation, 5 mM AMP-PNP was added to prevent additional phosphorylation and 10 μL aliquots were collected at 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 6, 10, 15, 20 and 25 min timepoints. Samples were then treated as previously described.[21]

NMR solution structure determination and relaxation experiments

Samples used for protein assignment of active EL_LovR contained 500 μM of ^{15}N-^{13}C labeled protein in 20 mM HEPES (pH 7.5), 3 mM NaN_3, 5 mM $BeCl_2$, 15 mM NaF, 10 mM

$MgCl_2$ and 10% D_2O. NMR data were collected at 25 °C on Varian Inova 600 and 800 MHz spectrometers and processed using NMRPipe[28] and NMRViewJ (One Moon Scientific)[29]. Chemical shift assignments for backbone and sidechain nuclei were obtained from 3D HNCACB, CBCA(CO)NH, HNCO, H(CCO)NH, C(CO)NH and HCCH-TOCSY experiments[30]. These chemical shift assignments focused on the major conformer present in each sample; while few low intensity peaks in several spectra suggested the presence of minor conformation(s), we were unable to assign these in a sequence-specific manner. All chemical shifts were deposited at BMRB with accession codes as follows: 25140 (apo-EL_LovR), 25141 (Mg^{2+} bound EL_LovR), 25137 (Mg^{2+}/BeF_3^- bound EL_LovR). Interproton distance restraints were obtained from 3D ^{15}N,^{13}C simultaneous edited NOESY spectra[31]. Hydrogen bond restraints were defined from backbone amide ^2H exchange protection factors[32] obtained from $^{15}N/^1H$ HSQC data collected on lyophilized U-^{15}N EL_LovR protein resuspended in 99.9% D_2O (25 °C, pH 7.5) to identify H-bond donors, while H-bond acceptors were identified from manual NOESY analysis. ϕ and ψ angle restraints were obtained from chemical shift analyses using TALOS-N[33]. Mg^{2+} and BeF_3^- ions were not explicitly included in the structure calculations, although both were present in the samples used for structural restraint measurements. Structures were calculated with automated NOESY spectra assignment using ARIA 2.2[34]. An ensemble of 20 conformers of the Mg^{2+}/BeF_3^- bound EL_LovR with excellent statistics (Table 1) were deposited into the PDB (accession code: 2MSW).

Titration with $MgCl_2$ and BeF_3^-

A 2 M stock solution of $MgCl_2$ was prepared and titrated into 500 μM U-^{15}N labeled EL_LovR in 10 mM Tris (pH 7.5), 50 mM NaCl, 1 mM DTT, 10% D_2O, at 25 °C.$^{15}N/^1H$ HSQC spectra were collected with 0–10 mM $MgCl_2$ present, processed as described above, and analyzed for peak intensity changes using NMRViewJ software[29]titration analysis. For BeF_3^- titration stock solutions of $BeCl_2$ and NaF were prepared at 1 M and 0.5 M respectively. $BeCl_2$ (ranging from 0–20 mM) was titrated to 500 μM protein in buffer containing 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT, 10% D_2O and 50 mM NaF. NaF (ranging from 0–50 mM) was titrated to 500 μM protein in buffer containing 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT, 10% D_2O and 1 mM $BeCl_2$. Data collection, processing and analysis were as described above.

Results

EL_LovR undergoes significant changes in global shape upon Mg^{2+} binding and phosphorylation

We started our studies by probing the effects of Mg^{2+} binding and phosphorylation on the oligomerization state of EL_LovR, as many RRs change among monomeric, dimeric and higher order states upon activation[12, 35, 36]. To determine the oligomeric state of EL_LovR, we used SEC-MALS analysis to obtain two independent types of data regarding protein mass and hydrodynamic behavior. We tested EL_LovR in three different conditions: apo (buffer only), metal-bound (buffer + 10 mM $MgCl_2$) and phosphorylated (buffer + 10 mM $MgCl_2$ + 10 mM carbamoyl phosphate). Under all conditions, light scattering established that EL_LovR is monomeric (~13 kDa); however, we observed substantial changes in

hydrodynamic behavior depending on the protein activation state (Figure 2). For the apo and Mg^{2+}-bound states, the ~14 mL elution volume is much smaller than expected from the EL_LovR sequence, consistent with a 53 kDa particle approximately four times the calculated molecular weight of EL_LovR. We attribute this to an expanded, monomeric EL_LovR given the monomeric molecular weight by MALS. In contrast, in comparable studies of phosphorylated EL_LovR (achieved by pre-incubation with carbamoyl phosphate), it eluted in two distinct peaks. One of these peaks corresponded to the elution volume of the apo and metal-bound state, which we assign to apo protein. However, the later-eluting peak (~17 mL) corresponded to the predicted molecular weight of monomeric EL_LovR, suggesting that the protein underwent a substantial compaction upon phosphorylation (with some fraction of the protein either not having been initially phosphorylated or having spontaneously dephosphorylated due to inherent phosphatase activity in this protein, *vide infra*). We note that the four-fold difference in apparent molecular weights attributable to the two SEC peaks corresponds to a 1.6-fold reduction in hydrodynamic radius; notably, this is less than the 1.8-fold difference expected between the folded and chemically-denatured states of a 125 residue protein like EL_LovR[37], suggesting that the apo form of EL_LovR is largely, but not completely, unfolded (analogous to some other proteins which undergo ligand-dependent folding[38]).

To gain more information on changes EL_LovR undergoes upon phosphorylation, we used limited proteolysis as a low-resolution structural probe. Using a well-established phosphomimic, beryllium fluoride (BeF_3^-), that non-covalently binds specifically to the phosphorylatable Asp residue located at the C-terminal end of the β3 strand[39, 40], to stably activate EL_LovR without issue of potential phosphatase activity, we incubated EL_LovR in varying conditions with trypsin. SDS-PAGE analysis of the resulting samples show that apo-and metal-bound EL_LovR undergoes rapid cleavage, completely converting from full length into a stable ~6 kDa fragment within 15 min (Figure 3a). ESI-MS analysis revealed the protease-resistant fragment to be a 5,851.1 Da segment of EL_LovR, corresponding to residues 8–62. In contrast, activated EL_LovR bound to Mg^{2+} and BeF_3^- remained markedly resistant to trypsin, with only minor proteolysis over time. These results suggest that EL_LovR undergoes significant conformational changes upon activation, consistent with an expanded, protease-accessible protein in the apo- and Mg^{2+}-bound forms folding into a compact, resistant structure upon phosphorylation.

EL_LovR becomes increasingly folded in activating conditions

To obtain residue-level analysis of the EL_LovR folding events implicated by SEC-MALS and limited proteolysis, we collected $^{15}N/^{1}H$ HSQC spectra of apo, Mg^{2+}-bound, Mg^{2+}/BeF_3^- bound samples (Figure 4). These data reveal significant improvements in ^{1}H chemical shift dispersion and homogeneity of peak intensity consistent with protein folding. In the apo form, most of the peaks remain clustered with poor amide proton dispersion in the center of the spectra, with some intense peaks present. These are all spectral signatures of a partially disordered protein with limited secondary and tertiary structure. We observed some improvement in peak dispersion and homogeneity upon addition of 10 mM $MgCl_2$, including shifting of some peaks outside of the central region. These results indicate conformational changes taking place that were otherwise undetectable with gel filtration or

limited proteolysis. To induce the fully active conformation, we incubated EL_LovR with 5 mM BeF_3^-, with the striking generation of outstanding peak dispersion consistent of a well-folded protein (Figure 4). Taken together, these data indicate that EL_LovR undergoes a folding event upon activation with BeF_3^-. Consistent with this, ^{15}N-edited NOESY data collected on EL_LovR in inactive and active conditions exhibited drastic differences (Figure S1, Supporting Information) with very few NOEs observed in the inactive (apo) state in contrast with many short and long range NOEs detected upon addition of Mg^{2+} and BeF_3^-.

We extended this analysis by completing backbone chemical shift assignments on the apo, Mg^2 and Mg^2/ BeF_3^- states of EL_LovR to obtain secondary structure analyses and chemical-shift derived order parameter information (S^2_{CS})[41](Figure 5). Using standard resonance approaches, we obtained backbone chemical shift assignments for 57%, 60% and 100% of residues in the major conformation present in each of these three conditions; the presence of a small number of low intensity peaks in several spectra suggest that some residues may adopt minor conformations as well, but we were unable to obtain sufficient data to assign these in a sequence-specific manner. For the apo and Mg^{2+}-bound EL_LovR species, intermediate chemical shift exchange also reduced peak intensities to the point of hampering chemical shift assignments in several regions of the major conformer (Figure 5b), including the trypsin-susceptible region identified in our limited proteolysis experiments. Those regions that were readily assigned showed significant disorder with fewer (and shorter) secondary structure elements than expected within REC domains (Figure 5a). In contrast, EL_LovR incubated with Mg^{2+} and BeF_3^- adopted a more structured conformation, with all REC domain secondary structure elements present (Figure 5b).

Further data supporting this ligand-dependent folding process was provided by comparisons of backbone amide order parameters extracted from chemical shifts (S^2_{CS})[41]. This parameter, which scales from 0 to 1 with increasing rigidity at any given backbone amide site, reveals that apo-EL_LovR is clearly unfolded, with few regions having S^2 ~0.8–0.9 values typically observed of well-folded protein. Addition of Mg^{2+} substantially elevates many of these values, particularly for residues in the N-terminal half of the domain. However, we still observed large variations in S^2 along the sequence, with low values (S^2~0.6) occurring in regions that usually adopt the stable α4 helix and β5 strand. Indeed, only when EL_LovR was incubated with both Mg^{2+} and BeF_3^- did we observe full ordering of the protein with S^2~0.8–0.9 values and all expected secondary structures present (Figure 5b).

To further assess the stability and dynamics of the fully-structured Mg^{2+} and BeF_3^- bound state of EL_LovR, we used a combination of ^2H exchange and ^{15}N relaxation measurements. While ^2H exchange measurements of the unfolded apo- and Mg^{2+}-bound species exchanged too quickly within the 20 min deadtime and duration of the first $^{15}N/^1H$ HSQC spectrum to accurately quantitate exchange rates, we measured ^2H protection factors of 10^5–10^6 for residues in secondary structure elements through most of the Mg^{2+}/BeF_3^- loaded state (Figure S2, Supporting Information). Intriguingly, this did not include residues in the α4-β5-α5 region that normally interacts with downstream effectors, as these exchanged too quickly to measure (Figure S2, Supporting Information). Complementary measurements of backbone dynamics via ^{15}N R_1 and R_2 relaxation rates paralleled these results (Figure S2b,

Supporting Information). Through most of the N-terminal portion of Mg^{2+}/BeF_3^- loaded EL_LovR, we observed uniform values of both relaxation parameters through secondary structure elements, consistent with minimal dynamics aside from overall tumbling. In contrast, we saw significant variability of both R_1 and R_2 values at several helical or strand sites in the α4-β5-α5 region, implicating greater flexibility in these regions on the ms (R_2) and ns-ps (R_1, R_2) timescales. Taken together, these data clearly establish stabilization of the EL_LovR REC domain fold by addition of Mg^{2+} and BeF_3^- ligands; the resulting fully-activated mimic state contains all of the expected structural elements of a REC domain while remaining dynamic in a region that is typically functionally important for dimerization and regulation of effector domains[6].

Solution structure of EL_LovR

Based on these observations, we sought to determine the NMR solution structure of activated EL_LovR to gain more insight into the structural features driving the ligand-induced folding process. Using standard solution NMR methods, we solved the structure of Mg^{2+} / BeF_3^- loaded EL_LovR with high precision (Figure 6a, Table 1), confirming that it adopts the canonical $(\beta\alpha)_5$ REC domain fold[42]. Examination of the structure shows that the phosphoaccepting D56 is located at the end of β3, adjacent to T83 on the subsequent β4 strand (Figure 6b,c). Furthermore, residues predicted to bind Mg^{2+} (E12, D13 and E14) are found in the loop region 1 (between β1 and α1), positioning them within ~5 Å of the oxygen atom of D56 (Figure 6b,c). Interestingly, V100 occupies the predicted position of the conserved Phe/Tyr residues involved in the Y-T coupling in REC domains[11, 43]. Additionally, the sidechain of a lysine residue (K103) also predicted to stabilize the active state of REC domains is oriented in close proximity to D56 (Figure 6b,c). All of these residues form a tight signaling network that runs from the N-terminal region of EL_LovR (D12, E13, E14, D56) all the way to the C-terminus (T83, K103) allowing the protein to properly signal upon phosphorylation[44, 45]. In summary, our biophysical characterization of EL_LovR conclusively establishes that it has minimal structure in the apo form, while addition of Mg^{2+} and BeF_3^- culminates in a properly-structured REC domain upon full activation.

Conserved residues involved in Mg^{2+} binding have a key role in structural changes of EL_LovR

With the presence of four closely-spaced acidic residues predicted for Mg^{2+} binding (E12, D13, E14 and D56), we hypothesized that repulsion among these negatively-charged residues in the EL_LovR active site would destabilize the protein in the absence of Mg^{2+} or another divalent cation, as previously observed for CheY[46]. To test this possibility, we mutated these four acidic residues in EL_LovR to alanine, and examined the resulting effects on protein structure in the absence and presence of Mg^{2+}. While one of the mutant proteins, E14A, appeared to be destabilized similarly as wild type based on ^1H chemical shift dispersion and heterogeneity of peak intensities in $^{15}N/^1H$ HSQC spectra (Figure 7), spectra of three other mutants (E12A, D13A and D56A) suggested these changes substantially stabilized the fold in the absence of Mg^{2+}. We interpret these data to indicate a greater role for D13 and D56 (and to a lesser degree, E12) than E14 in electrostatic destabilization of the native fold. Upon addition of $MgCl_2$, we observed more limited

chemical shift changes for each of the four mutant proteins than wildtype (Figures 4,7), demonstrating the importance of all four residues in establishing the Mg^{2+} binding site (Figure 7). Notably, incubating EL_LovR D56A with both $MgCl_2$ and BeF_3^- also showed no changes in the NMR spectra from the apo form of this protein, supporting the assignment of D56 as the phosphoaccepting Asp residue EL_LovR (Figure S3, Supporting Information). To assess the effects of these mutations more globally on EL_LovR structure, we conducted SEC-MALS analyses of D13A, D56A variants and a D13A/D56A double mutant, observing shifts to later elution volumes on Superdex S200 gel filtration while the protein remained monomeric (Figure S4, Supporting Information). These shifts, reminiscent of changes observed upon phosphorylation in wildtype EL_LovR, are consistent with increased stabilization of the EL_LovR REC domain fold, underscoring the importance of the cluster of negatively-charged residues in destabilizing EL_LovR without ligand.

EL_LovR has a short-lived phosphorylated state

Genetic analyses of the *C. crescentus* LovR homolog of EL_LovR strongly suggest that it functions as a phosphate sink[47], by depleting the phosphoryl group from the HK and quickly losing it to the solvent by hydrolysis, in stress response pathways in that organism[23, 24], leading us to hypothesize a similar role for EL_LovR. We tested this possibility with a series of *in vitro* dephosphorylation assays, using the EL368 kinase to initially generate EL_LovR~P. Once this was completed, we added a non-hydrolyzable ATP analogue AMP-PNP to stop further incorporation of phosphate into the RR and allow solely dephosphorylation to occur. We observed the expected time-dependence of residual EL_LovR~P after AMP-PNP addition, with EL_LovR completely dephosphorylated after approximately 25 min (Figure 8). Fitting the time dependence of this process to a single exponential decay, we calculated the rate constant of dephosphorylation to be k= 0.27 min^{-1} (approximately a 2.5 min half-life) roughly comparable to CheY, a well-known phosphate sink[48]. Another parallel between CheY and EL_LovR is the presence of an asparagine residue nearby the phosphoaccepting aspartate; within CheY, this residue promotes the dephosphorylation reaction by coordinating a water molecule near the active site[48]. Dephosphorylation assays of EL_LovR (N58K), which contains a mutation in this asparagine, shows that this protein has a greatly slowed dephosphorylation rate (half-life >> 15 min) suggesting a comparable role for this residue as in CheY. Coupled with our prior demonstration of EL_LovR indirectly inducing the dephosphorylation of the EL_PhyR response regulator *in vitro* by promoting the phosphatase activity of the HK[21], these observations suggest that EL_LovR functions as a phosphate sink within *E. litoralis* stress response pathways.

Discussion

Given their essential roles within many bacterial signaling pathways and ease of study, REC domains have served as model systems for allosteric signaling proteins. As such, understanding how these molecular switches function has been of widespread interest for many years, with a generally-agree model arising from crystallographic and NMR studies of phosphorylation-dependent changes within a broad group of REC domains[10, 13, 19, 20, 49–51]. This model is centered on a phosphorylation-dependent switch between two well-folded

states, triggered by interactions between the newly-incorporated phosphoryl group and a nearby Ser/Thr hydroxyl group from a residue in the β4 strand. Subsequently, an aromatic residue in the β5 strand rotates to fill the pocket previously occupied by the Ser/Thr residue. This so-called "Y-T coupling" process[43] is believed to be commonly (but perhaps not universally[52]) involved in propagating structural changes from the active site to the α4-β5-α5 interface, a region commonly used by REC domains to regulate function and protein-protein interactions[53]. Viewed at the domain level, solution NMR studies indicate that unphosphorylated REC domains sample well-folded inactive and active conformations[19, 54–56] in a pre-existing equilibrium, allowing phosphorylation to activate the protein by simply shifting the bias between states.

In contrast with most other RR proteins, our biophysical analyses of EL_LovR suggest that it differs from this model by having phosphorylation control protein folding, rather than simply switching an equilibrium among well-folded states. From multiple independent viewpoints, EL_LovR undergoes a progressive stabilization of the REC domain fold as its Mg^{2+} and phosphoryl ligands are added. Only once fully activated by both of these groups does EL_LovR adopt the conserved REC domain fold, as demonstrated by our high-resolution structure. However, even in this state, EL_LovR still retains some unusual flexibility in the α4-β5-α5 region, as supported by hydrogen exchange and ^{15}N relaxation measurements. This is reminiscent of another recently described SDRR, Sma0114[14, 18], which contains a disordered α4 helix in both the inactive and active states.

While our data indicate that both Mg^{2+} and BeF_3^- appear to be important for fully folding EL_LovR, it is worth noting that some regions of the protein order simply upon the addition of divalent metal cations. Key among these is the β1-α1-β2-α2-β3 region, which appears to fold and become ordered in 10 mM $MgCl_2$ as demonstrated by backbone NMR chemical shifts. This region contains the residues that coordinate Mg^{2+} (E12, D13 and E14) near the phosphoaccepting D56 residue. In the absence of divalent metals, electrostatic repulsion among these acidic residues appears to destabilize EL_LovR, as borne out by the greater structural character of the D13A, D56A and D13A/D56A proteins by SEC-MALS and solution NMR. Furthermore, despite the close proximity among those negatively charged residues (E12, D13, E14), our data suggests that the first two of these residues at the end of the β1 strand (E12, D13) play a major role in the metal binding, which is highly correlated with their degree of conservation[4, 51]. Within the wildtype protein, addition of $MgCl_2$ is sufficient to overcome this repulsion and order the N-terminal region of the protein. We speculate that this stabilization is important for priming EL_LovR for phosphorylation by its cognate histidine kinases, because α1 helix residues are known to make contacts with HKs (and hence dictate HK/RR specificity) among REC domains[57]. Similar electrostatic effects on the stability of protein structure have been observed in a range of different systems, including the Sic1 cyclin dependent kinase inhibitor, the "metamorphic" chemokine lymphotactin to SH3 protein/protein interaction domains[38, 58, 59]. In each case, clusters of positively- or negatively-charged residues with accompanying electrostatic repulsion destabilize the protein structure; such repulsion can be counteracted by neutralization through mutagenesis, binding oppositely-charged ligands, or increasing the ionic strength of the surrounding environment.

Given that Mg^{2+} levels in the native marine environment of *E. litoralis* are approximately 50 mM[60], we anticipate that the partially-folded Mg^{2+} bound state – ready to interact with an HK, but not yet fully structured – represents the "inactive" state of EL_LovR in the cell in marked contrast to the fully-ordered inactive states of NtrC, Spo0F and CheY proteins[16, 17, 50]. We note that a substantial number of isolated REC domain structures have been solved by X-ray diffraction or solution NMR without any divalent cations present (*e.g.* PhoB, NtrC, PhoP[16, 61, 62]) but are all fully folded. These data suggest some diversity among REC domains in the degree of folding they adopt between functional states, with the potential for sequence/structural variants to substantially tune these features. For instance, REC domains with substitutions in the active site (close to the phosphoaccepting Asp) and conserved regions (for example, the β4-α4-β5 interface) are prime candidates to have unusual characteristics that may diverge from classical RRs. Aligning the EL_LovR sequence to those of related REC domains (*C. crescentus* LovR and Sma0114) and well-studied REC domains (CheY, Spo0F and NtrC), we observe great diversity in the conserved β4-α4-β5 interface (Figure S5, Supporting Information). Within EL_LovR, a clear candidate for such a residue is Pro 99, located within the β5 region. Secondary structure analyses (using STRIDE[63]) of EL_LovR and two well-studied SDRRs (CheY and Spo0F[17, 64]) reveal that EL_LovR has a shorter β5-strand than other SDRRs, suggesting that Pro 99 destabilizes the N-terminal half of this structural element and potentially more broadly to the domain fold overall. Additionally, at the end of β4 and α4, several residues (His82, Ile 96 and Asp 97) notably differ from those found in other well-known RRs. These three His/Ile/Asp residues are typically occupied by an Ala/Gly/Ala triple; these residues are found in loops between β4-α4-β5, allowing the formation of these secondary structure elements. We speculate that substitutions of the small Gly and Ala residues with larger (and potentially charged) residues found in EL_LovR result in clashes with other sites in the protein that destabilize the α4 helix and adjacent regions. These and other non-conservative alterations may play a role in the conformational changes EL_LovR undergoes upon phosphorylation.

Turning from structural aspects of EL_LovR to functional characteristics, several lines of evidence suggest that it serves as a phosphatase to downregulate signaling through a general stress pathway in *E. litoralis*. While this organism has remained genetically intractable in our hands, studies of the highly-related (50% sequence identity) LovR protein in the *Caulobacter crescentus* model organism supports this assertion. When *C. crescentus* is under osmotic stress, the PhyK histidine kinase is activated and phosphorylates its cognate RR (PhyR)[23, 65], activating the expression of σ^T-regulated genes including the *lovR* locus itself. Evidence suggests that LovR in turn acts as a phosphatase to PhyR~P, serving as a negative feedback loop of σ^T driven transcription[23]. Consistent with this role, we observed marked *in vitro* phosphatase activity of EL_LovR, with a half-life of phospho-EL_LovR of approximately 2 min. We note that some RRs are capable of catalyzing the dephosphorylation reaction on their own while others rely on accessory proteins to fully accelerate this reaction[66]. In these latter cases, an asparagine or glutamine residue is often used to coordinate a water molecule in the active site, promoting hydrolysis of the phospho-Asp[66]. Closer examination of the sequence in the EL_LovR active site shows that an asparagine residue (N58) is proximal to D56; we suggest this residue may serve to coordinate water and catalyze the dephosphorylation reaction of EL_LovR~P, as proposed

for CheY[48]. More broadly, bioinformatics analysis shows that *E. litoralis* contains genes homologous to the *C. crescentus* system[21], suggesting conservation of this signaling pathway among these alpha-proteobacteria. Initial *in vitro* experiments show that co-incubation of EL_LovR and its EL368 HK induces dephosphorylation of EL_PhyR~P, indicating that these proteins act similarly to their *C. crescentus* counterparts[21]. Critically, a role for EL_LovR as a phosphatase does not require that it interact directly with any downstream effector; such a role might be complicated by the partially destabilized α4-β5-α5 surface that usually provides the surface for such interactions with other REC domains. However, as we detail above, the EL_LovR α1 helix is folded by the presence of Mg^{2+}, ordering the critical secondary structure element needed to interact with upstream kinases for the initial phosphorylation event to occur[57].

Taken together, these studies provide evidence of a SDRR that has structural features that differ from characterized RRs. In particular, the ability of EL_LovR to undergo a folded transition is an example of a new type of conformational changes exhibited by a REC domain, which could have implications on the regulatory modes for these diverse signaling domains.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This work was supported by NIH grants R01 GM106239 (K.H.G.) and T32 AI007520 (supporting V.J.O.).

References

1. Capra EJ, Laub MT. Evolution of two-component signal transduction systems. Annu. Rev. Microbiol. 2012; 66:325–347. [PubMed: 22746333]
2. Stock AM, Robinson VL, Goudreau PN. Two-component signal transduction. Annu. Rev. Biochem. 2000; 69:183–215. [PubMed: 10966457]
3. Bourret RB, Silversmith RE. Two-component signal transduction. Curr. Opin. Microbiol. 2010; 13:113–115. [PubMed: 20219418]
4. Lukat GS, Stock AM, Stock JB. Divalent metal ion binding to the CheY protein and its significance to phosphotransfer in bacterial chemotaxis. Biochemistry. 1990; 29:5436–5442. [PubMed: 2201404]
5. Jenal U, Galperin MY. Single domain response regulators: molecular switches with emerging roles in cell organization and dynamics. Curr. Opin. Microbiol. 2009; 12:152–160. [PubMed: 19246239]
6. Gao R, Mack TR, Stock AM. Bacterial response regulators: versatile regulatory strategies from common domains. Trends Biochem. Sci. 2007; 32:225–234. [PubMed: 17433693]
7. Bren A, Eisenbach M. The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY. J. Mol. Biol. 1998; 278:507–514. [PubMed: 9600834]
8. Jacobs C, Hung D, Shapiro L. Dynamic localization of a cytoplasmic signal transduction response regulator controls morphogenesis during the Caulobacter cell cycle. Proc Natl Acad Sci U S A. 2001; 98:4095–4100. [PubMed: 11274434]
9. Hung DY, Shapiro L. A signal transduction protein cues proteolytic events critical to Caulobacter cell cycle progression. Proc Natl Acad Sci U S A. 2002; 99:13160–13165. [PubMed: 12237413]

10. Kern D, Volkman BF, Luginbuhl P, Nohaile MJ, Kustu S, Wemmer DE. Structure of a transiently phosphorylated switch in bacterial signal transduction. Nature. 1999; 402:894–898. [PubMed: 10622255]
11. Cho HS, Lee SY, Yan D, Pan X, Parkinson JS, Kustu S, Wemmer DE, Pelton JG. NMR structure of activated CheY. J. Mol. Biol. 2000; 297:543–551. [PubMed: 10731410]
12. Lee SY, Cho HS, Pelton JG, Yan D, Berry EA, Wemmer DE. Crystal structure of activated CheY. Comparison with other activated receiver domains. J. Biol. Chem. 2001; 276:16425–16431. [PubMed: 11279165]
13. Gardino AK, Volkman BF, Cho HS, Lee SY, Wemmer DE, Kern D. The NMR solution structure of BeF(3)(−)-activated Spo0F reveals the conformational switch in a phosphorelay system. J. Mol. Biol. 2003; 331:245–254. [PubMed: 12875849]
14. Sheftic SR, White E, Gage DJ, Alexandrescu AT. NMR Structure of the HWE Kinase Associated Response Regulator Sma0114 in Its Activated State. Biochemistry. 2014; 53:311–322. [PubMed: 24364624]
15. Bellsolell L, Prieto J, Serrano L, Coll M. Magnesium binding to the bacterial chemotaxis protein CheY results in large conformational changes involving its functional surface. J. Mol. Biol. 1994; 238:489–495. [PubMed: 8176739]
16. Volkman BF, Nohaile MJ, Amy NK, Kustu S, Wemmer DE. Three-dimensional solution structure of the N-terminal receiver domain of NTRC. Biochemistry. 1995; 34:1413–1424. [PubMed: 7827089]
17. Feher VA, Zapf JW, Hoch JA, Whiteley JM, McIntosh LP, Rance M, Skelton NJ, Dahlquist FW, Cavanagh J. High-resolution NMR structure and backbone dynamics of the Bacillus subtilis response regulator, Spo0F: implications for phosphorylation and molecular recognition. Biochemistry. 1997; 36:10015–10025. [PubMed: 9254596]
18. Sheftic SR, Garcia PP, Robinson VL, Gage DJ, Alexandrescu AT. NMR assignments for the Sinorhizobium meliloti response regulator Sma0114. Biomol NMR Assign. 2011; 5:55–58. [PubMed: 20936511]
19. Volkman BF, Lipson D, Wemmer DE, Kern D. Two-state allosteric behavior in a single-domain signaling protein. Science. 2001; 291:2429–2433. [PubMed: 11264542]
20. McDonald LR, Boyer JA, Lee AL. Segmental motions, not a two-state concerted switch, underlie allostery in CheY. Structure. 2012; 20:1363–1373. [PubMed: 22727815]
21. Correa F, Ko WH, Ocasio V, Bogomolni RA, Gardner KH. Blue light regulated two-component systems: enzymatic and functional analyses of light-oxygen-voltage (LOV)-histidine kinases and downstream response regulators. Biochemistry. 2013; 52:4656–4666. [PubMed: 23806044]
22. Rivera-Cancel G, Ko WH, Tomchick DR, Correa F, Gardner KH. Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. Proc Natl Acad Sci U S A. 2014; 111:17839–17844. [PubMed: 25468971]
23. Foreman R, Fiebig A, Crosson S. The LovK-LovR two-component system is a regulator of the general stress pathway in Caulobacter crescentus. J. Bacteriol. 2012; 194:3038–3049. [PubMed: 22408156]
24. Purcell EB, Siegal-Gaskins D, Rawling DC, Fiebig A, Crosson S. A photosensory two-component system regulates bacterial cell attachment. Proc Natl Acad Sci U S A. 2007; 104:18241–18246. [PubMed: 17986614]
25. Oh HM, Giovannoni SJ, Ferriera S, Johnson J, Cho JC. Complete genome sequence of Erythrobacter litoralis HTCC2594. J. Bacteriol. 2009; 191:2419–2420. [PubMed: 19168610]
26. Sheffield P, Garrard S, Derewenda Z. Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. Protein Expr Purif. 1999; 15:34–39. [PubMed: 10024467]
27. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 1976; 72:248–254. [PubMed: 942051]
28. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR. 1995; 6:277–293. [PubMed: 8520220]

29. Johnson BA, Blevins RA. NMR View: A computer program for the visualization and analysis of NMR data. J. Biomol. NMR. 1994; 4:603–614. [PubMed: 22911360]
30. Cavanagh, J.; Fairbrother, WJ.; Palmer, AG.; Rance, M.; Skelton, NJ. Protein NMR Spectroscopy: Principles and Practice. San Diego, CA: Elsevier; 2007.
31. Pascal SM, Muhandiram DR, Yamazaki T, Forman-Kay JD, Kay LE. Simultaneous Acquisition of N-15-Edited and C-13-Edited Noe Spectra of Proteins Dissolved in H2O. J. Magn. Reson., Ser B. 1994; 103:197–201.
32. Bai Y, Milne JS, Mayne L, Englander SW. Primary structure effects on peptide group hydrogen exchange. Proteins. 1993; 17:75–86. [PubMed: 8234246]
33. Shen Y, Bax A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J. Biomol. NMR. 2013; 56:227–241. [PubMed: 23728592]
34. Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE, Nilges M. ARIA2: automated NOE assignment and data integration in NMR structure calculation. Bioinformatics. 2007; 23:381–382. [PubMed: 17121777]
35. Wyman C, Rombel I, North AK, Bustamante C, Kustu S. Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. Science. 1997; 275:1658–1661. [PubMed: 9054362]
36. Creager-Allen RL, Silversmith RE, Bourret RB. A link between dimerization and autophosphorylation of the response regulator PhoB. J. Biol. Chem. 2013; 288:21755–21769. [PubMed: 23760278]
37. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. Biochemistry. 1999; 38:16424–16431. [PubMed: 10600103]
38. Zhang O, Forman-Kay JD. Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. Biochemistry. 1995; 34:6784–6794. [PubMed: 7756310]
39. Yan D, Cho HS, Hastings CA, Igo MM, Lee SY, Pelton JG, Stewart V, Wemmer DE, Kustu S. Beryllofluoride mimics phosphorylation of NtrC and other bacterial response regulators. Proc Natl Acad Sci U S A. 1999; 96:14789–14794. [PubMed: 10611291]
40. Lee SY, Cho HS, Pelton JG, Yan D, Henderson RK, King DS, Huang L, Kustu S, Berry EA, Wemmer DE. Crystal structure of an activated response regulator bound to its target. Nat. Struct. Biol. 2001; 8:52–56. [PubMed: 11135671]
41. Berjanskii MV, Wishart DS. A simple method to predict protein flexibility using secondary chemical shifts. J. Am. Chem. Soc. 2005; 127:14970–14971. [PubMed: 16248604]
42. Gao R, Stock AM. Biological insights from structures of two-component proteins. Annu. Rev. Microbiol. 2009; 63:133–154. [PubMed: 19575571]
43. Stock AM, Guhaniyogi J. A new perspective on response regulator activation. J. Bacteriol. 2006; 188:7328–7330. [PubMed: 17050920]
44. Lukat GS, Lee BH, Mottonen JM, Stock AM, Stock JB. Roles of the highly conserved aspartate and lysine residues in the response regulator of bacterial chemotaxis. J. Biol. Chem. 1991; 266:8348–8354. [PubMed: 1902474]
45. Appleby JL, Bourret RB. Proposed signal transduction role for conserved CheY residue Thr87, a member of the response regulator active-site quintet. J. Bacteriol. 1998; 180:3563–3569. [PubMed: 9657998]
46. Smith JG, Latiolais JA, Guanga GP, Citineni S, Silversmith RE, Bourret RB. Investigation of the role of electrostatic charge in activation of the Escherichia coli response regulator CheY. J. Bacteriol. 2003; 185:6385–6391. [PubMed: 14563873]
47. Szurmant H, Bunn MW, Cho SH, Ordal GW. Ligand-induced conformational changes in the Bacillus subtilis chemoreceptor McpB determined by disulfide crosslinking in vivo. J. Mol. Biol. 2004; 344:919–928. [PubMed: 15544802]
48. Thomas SA, Brewster JA, Bourret RB. Two variable active site residues modulate response regulator phosphoryl group stability. Mol. Microbiol. 2008; 69:453–465. [PubMed: 18557815]
49. McDonald LR, Whitley MJ, Boyer JA, Lee AL. Colocalization of fast and slow timescale dynamics in the allosteric signaling protein CheY. J. Mol. Biol. 2013; 425:2372–2381. [PubMed: 23648838]

50. Volz K, Matsumura P. Crystal structure of Escherichia coli CheY refined at 1.7-A resolution. J. Biol. Chem. 1991; 266:15511–15519. [PubMed: 1869568]
51. Bourret RB. Receiver domain structure and function in response regulator proteins. Curr. Opin. Microbiol. 2010; 13:142–149. [PubMed: 20211578]
52. Villali J, Pontiggia F, Clarkson MW, Hagan MF, Kern D. Evidence Against the "Y-T Coupling" Mechanism of Activation in the Response Regulator NtrC. J. Mol. Biol. 2014; 426:1554–1567. [PubMed: 24406745]
53. Gao R, Stock AM. Molecular strategies for phosphorylation-mediated regulation of response regulator activity. Curr. Opin. Microbiol. 2010; 13:160–167. [PubMed: 20080056]
54. Birck C, Mourey L, Gouet P, Fabry B, Schumacher J, Rousseau P, Kahn D, Samama JP. Conformational changes induced by phosphorylation of the FixJ receiver domain. Structure. 1999; 7:1505–1515. [PubMed: 10647181]
55. Feher VA, Cavanagh J. Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. Nature. 1999; 400:289–293. [PubMed: 10421374]
56. Gardino AK, Villali J, Kivenson A, Lei M, Liu CF, Steindel P, Eisenmesser EZ, Labeikovsky W, Wolf-Watz M, Clarkson MW, Kern D. Transient non-native hydrogen bonds promote activation of a signaling protein. Cell. 2009; 139:1109–1118. [PubMed: 20005804]
57. Podgornaia AI, Laub MT. Determinants of specificity in two-component signal transduction. Curr. Opin. Microbiol. 2013; 16:156–162. [PubMed: 23352354]
58. Tyler RC, Wieting JC, Peterson FC, Volkman BF. Electrostatic optimization of the conformational energy landscape in a metamorphic protein. Biochemistry. 2012; 51:9067–9075. [PubMed: 23102260]
59. Liu B, Chia D, Csizmok V, Farber P, Forman-Kay JD, Gradinaru CC. The effect of intrachain electrostatic repulsion on conformational disorder and dynamics of the Sic1 protein. J. Phys. Chem. B. 2014; 118:4088–4097. [PubMed: 24673507]
60. Elderfield H, Schultz A. Mid-ocean ridge hydrothermal fluxes and the chemical composition of the ocean. Annu. Rev. Earth Planet. Sci. 1996; 24:191–224.
61. Sola M, Gomis-Ruth FX, Serrano L, Gonzalez A, Coll M. Three-dimensional crystal structure of the transcription factor PhoB receiver domain. J. Mol. Biol. 1999; 285:675–687. [PubMed: 9878437]
62. Bachhawat P, Stock AM. Crystal structures of the receiver domain of the response regulator PhoP from Escherichia coli in the absence and presence of the phosphoryl analog beryllofluoride. J. Bacteriol. 2007; 189:5987–5995. [PubMed: 17545283]
63. Heinig M, Frishman D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. 2004; 32:W500–W502. [PubMed: 15215436]
64. Stock AM, Martinez-Hackert E, Rasmussen BF, West AH, Stock JB, Ringe D, Petsko GA. Structure of the Mg(2+)-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. Biochemistry. 1993; 32:13375–13380. [PubMed: 8257674]
65. Herrou J, Foreman R, Fiebig A, Crosson S. A structural model of anti-anti-sigma inhibition by a two-component receiver domain: the PhyR stress response regulator. Mol. Microbiol. 2010; 78:290–304. [PubMed: 20735776]
66. Silversmith RE. Auxiliary phosphatases in two-component signal transduction. Curr. Opin. Microbiol. 2010; 13:177–183. [PubMed: 20133180]

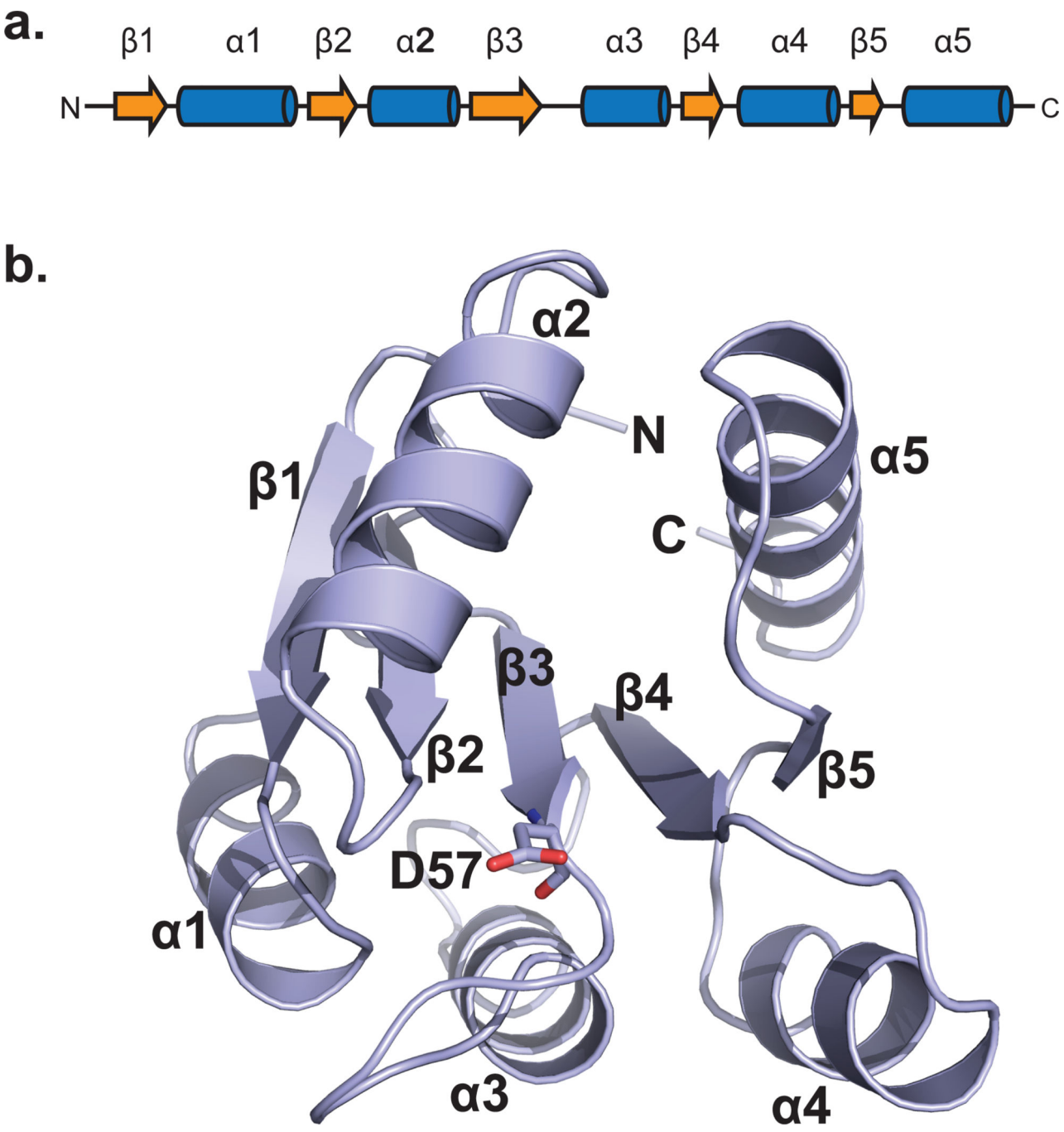

Figure 1. REC domain secondary and tertiary structure

a). Standard secondary structure elements and nomenclature in a typical REC domain. b). Crystal structure of the chemotaxis SDRR protein $CheY^{64}$, showing the canonical $(\beta\alpha)_5$ fold of REC domains.

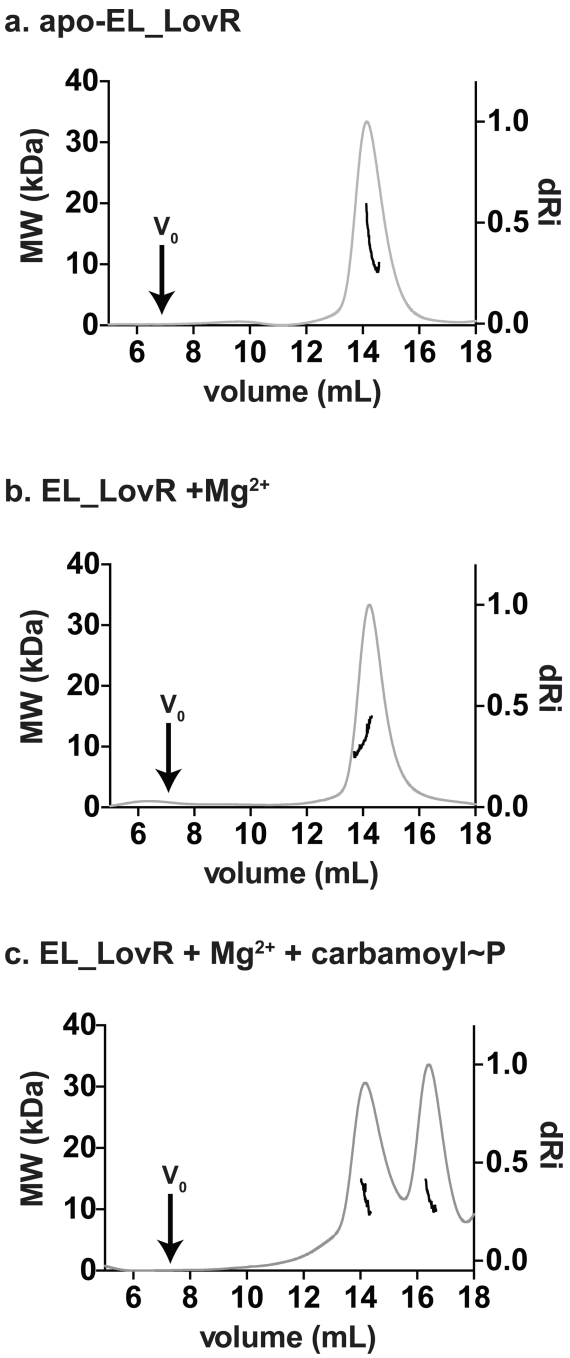

Figure 2. EL_LovR undergoes global conformational changes upon phosphorylation

a). Superdex 200 10/300 SEC-MALS data indicate that apo EL_LovR eluted with a 14 mL retention volume (differential refractive index, dRi; gray). The molecular weight obtained from MALS (black) indicates that apo-EL_LovR is a monomer at ~13 kDa while the apparent molecular weight based on elution volume is ~53 kDa. b). Comparable to the apo-protein, Mg^{2+}-bound EL_LovR eluted at 14 mL and with a monomeric molecular weight, suggesting that Mg^{2+} did not globally change the apo- state structure. c) Phosphorylated, Mg^{2+}-bound EL_LovR eluted with in two distinct peaks, one of which resembles the apo-

state while the other eluted with a later, 17 mL volume. We interpret the appearance of two peaks as indicating the sample containing both unphosphorylated and phosphorylated forms of EL_LovR due to the rapid phosphatase activity of EL_LovR~P. These data suggest that the overall shape of the molecule changed from an elongated state to a compact state while remaining monomeric ~13 kDa. The void volume (V_0) of this column is approximately 7 mL as indicated on all chromatograms.

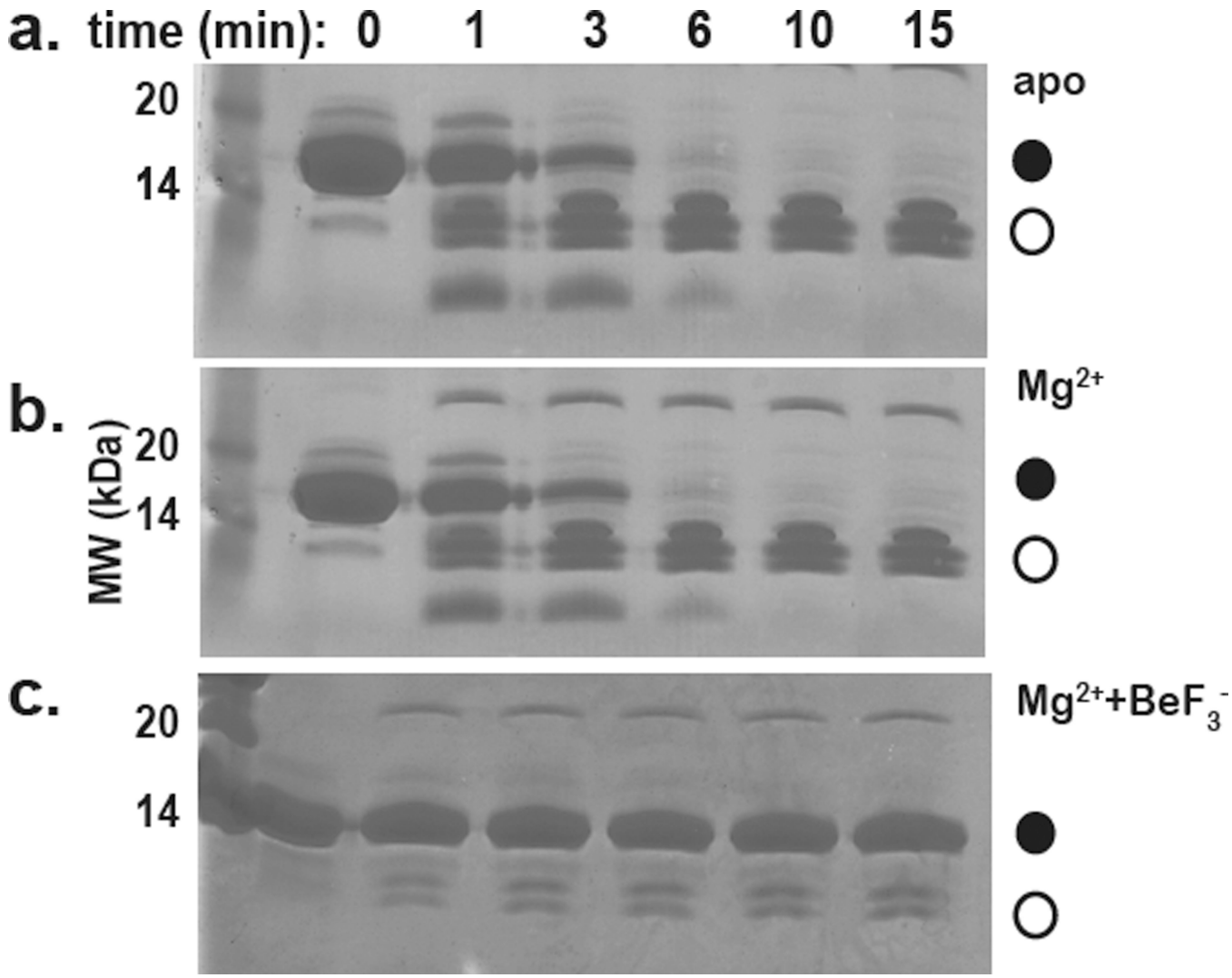

Figure 3. EL_LovR~P undergoes phosphorylation-dependent conformational changes to become more resistant to trypsin digestion

Limited proteolysis of EL_LovR with trypsin in a 1:90 w:w ratio (EL_LovR:trypsin) shows different patterns of cleavage of EL_LovR when phosphorylated. a). Apo-EL_LovR is rapidly cleaved by trypsin from its ~13 kDa full length form (closed circles) into a protease-resistant 5.8 kDa fragment (open circles), identified by ESI-MS to correspond to residues 8–62. b). Mg^{2+} -bound EL_LovR shows a very similar time dependence of trypsin digestion as apo- protein, c). Fully-activated EL_LovR bound to both Mg^{2+} and BeF_3^- was not degraded by trypsin over the same 15 min timecourse as apo- and Mg^{2+} -bound protein, consistent with substantial conformational changes upon activation.

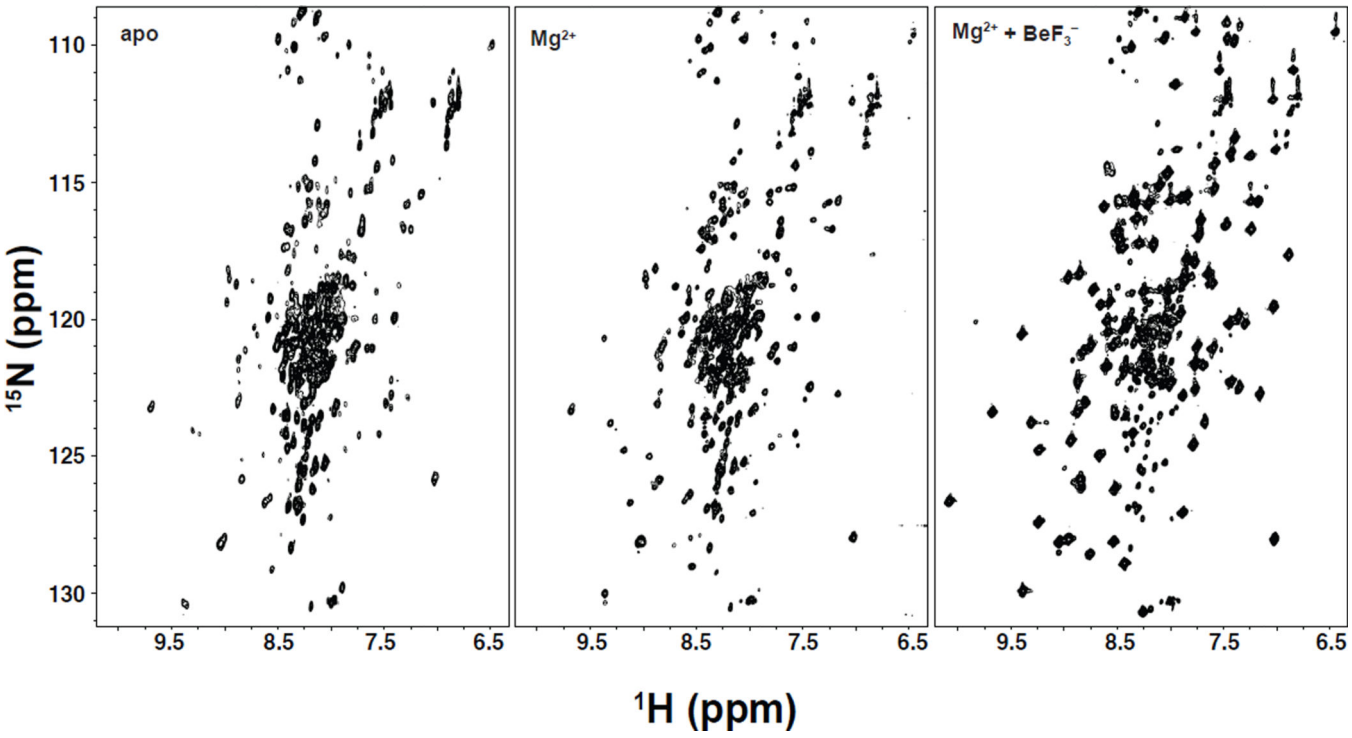

Figure 4. EL_LovR REC fold progressively becomes more stable with binding to Mg^{2+} and $BeF_3^{-15}N/^1$

H HSQC spectra were acquired of EL_LovR with varying ligands provided as indicated: Mg^{2+} = 10 mM $MgCl_2$; BeF_3^- = 5 mM $BeCl_2$ and 15 mM NaF to generate saturating BeF_3^-. Apo EL_LovR spectra are consistent with mostly disordered proteins, with relatively poor amide proton chemical shift dispersion and heterogeneous peak intensities. Addition of Mg^{2+} improves dispersion and reduces heterogeneity, as well as inducing some peak shifts; all of these signs suggest some degree of ligand-induced stabilization. By adding both Mg^{2+} and BeF_3^- to generate a fully activated state, we observed further improvement in peak dispersion and intensity, consistent with a fully-folded protein.

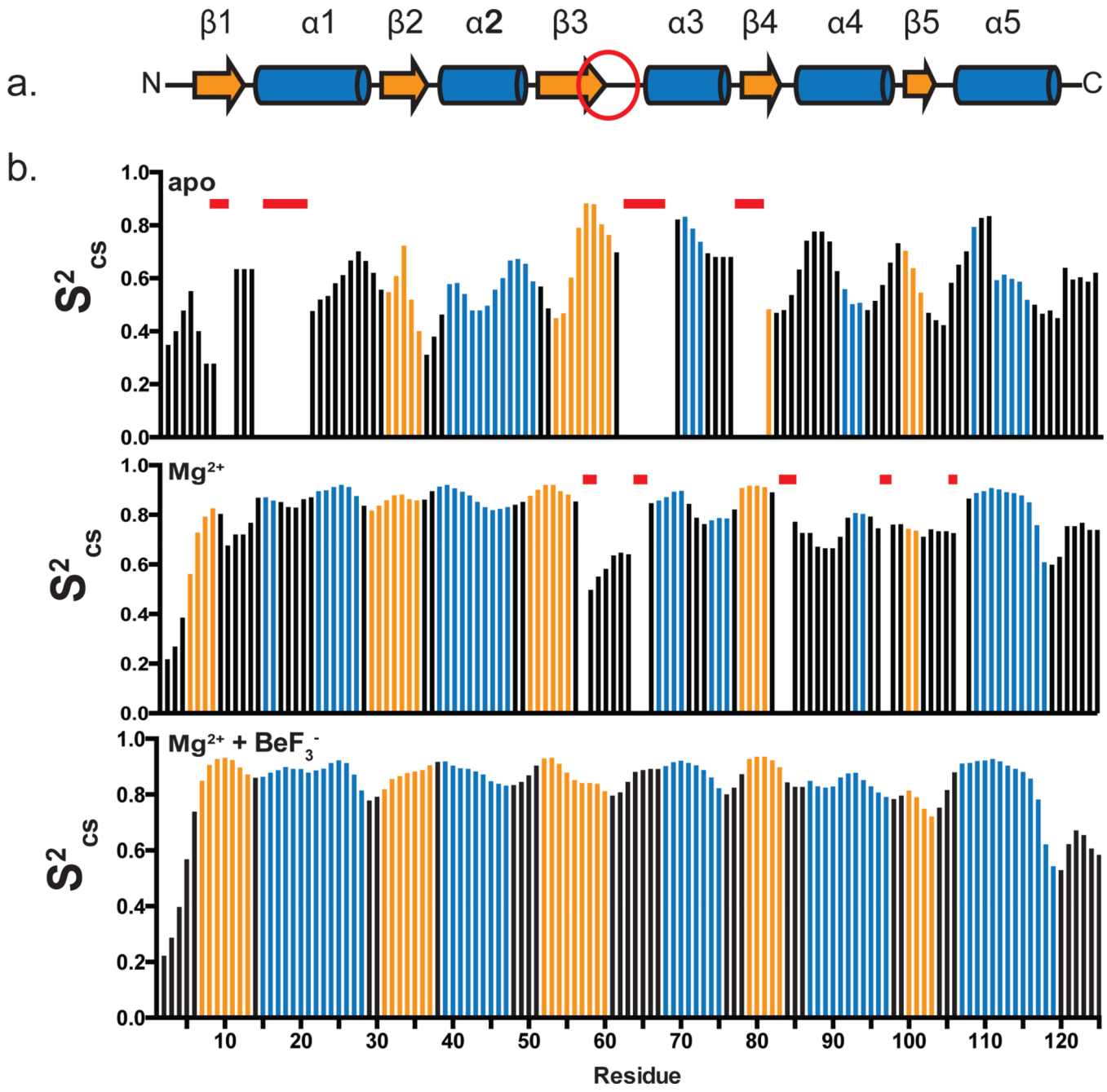

Figure 5. Ligand binding to EL_LovR induces the formation of stable secondary structure

a) Predicted secondary structure elements of EL_LovR based on REC domain structures; red circle indicates the location of the phosphoaccepting D56 residue. b). Analyses of EL_LovR backbone chemical shifts with TALOS-N[33] to identify secondary structure elements and characterize backbone mobility. Analysis of apo shifts revealed several secondary structure elements (yellow: helices; blue: strands, black: no regular secondary structure), although truncated compared to expected boundaries. S^2_{CS} analyses of the same chemical shifts predict significant dynamics on the ps-ns timescale throughout much of the protein. Red bars indicate residues without chemical shift assignments due to intermediate chemical

exchange. Comparable analysis of the Mg^{2+}-bound state showed additional secondary structure elements, particularly in the N-terminal half of the protein and increased S^2_{CS} values, consistent with ligand-induced order. Finally, addition of both Mg^{2+} and BeF_3^- leads to EL_LovR adopting the canonical $(\beta\alpha)_5$ topology of a REC domain, and an overall increase of S^2_{CS} values to 0.8–0.9 for most of the protein, indicating limited ns-ps timescale dynamics.

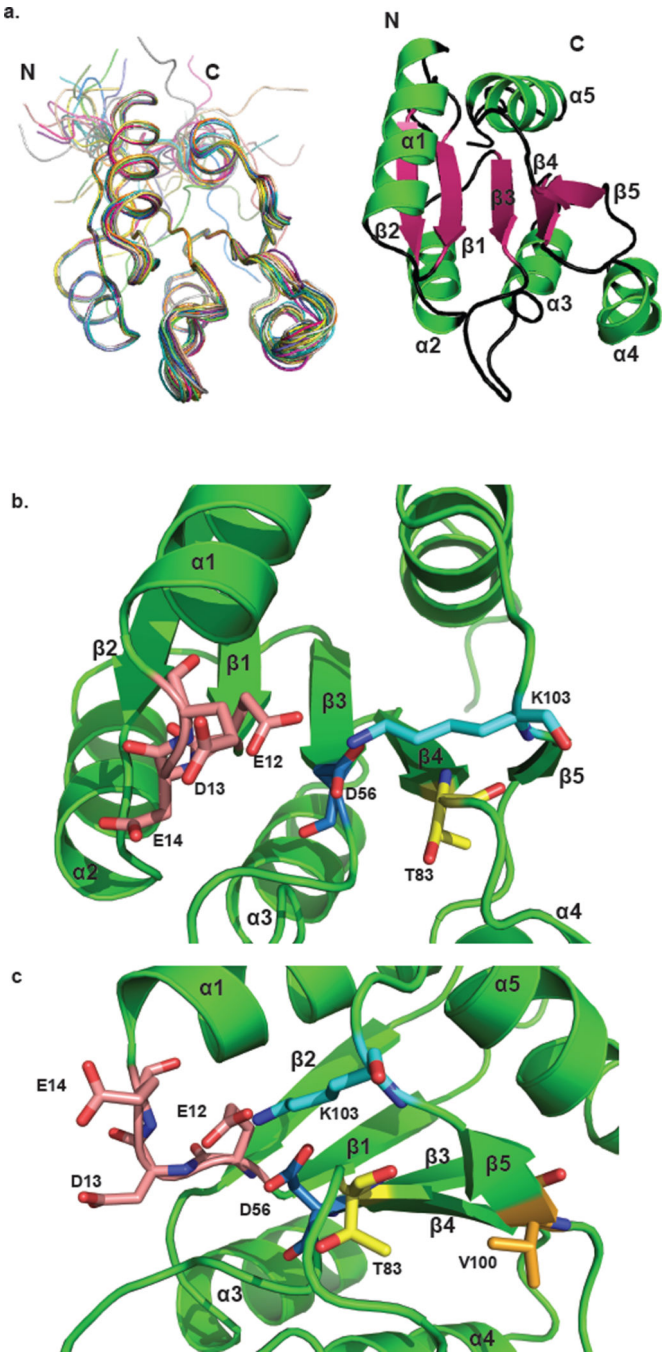

Figure 6. NMR solution structure of EL_LovR + Mg^{2+} + BeF_3

a). Ensemble of twenty solution structure models of activated EL_LovR, as calculated by ARIA 2.2[34] and ribbon diagram of lowest-energy structure. b). The EL_LovR active site is diagrammed, with key residues highlighted. Acidic residues involved in coordinating Mg^{2+} (E12, D13, E14) are indicated in pink; the phosphoaccepting D56 (blue) is positioned at the end of β3 in close proximity to T83 (yellow), with the sidechain of the conserved K103 residue (cyan) oriented towards the active site to stabilize the active state. c). The α4-(β5-α5 region of EL_LovR, showing the locations of the D56 – T84 – V100 signaling pathway,

where V100 occupies the position of the conserved tyrosine involved in “Y-T coupling” found in many REC domains[11, 52]

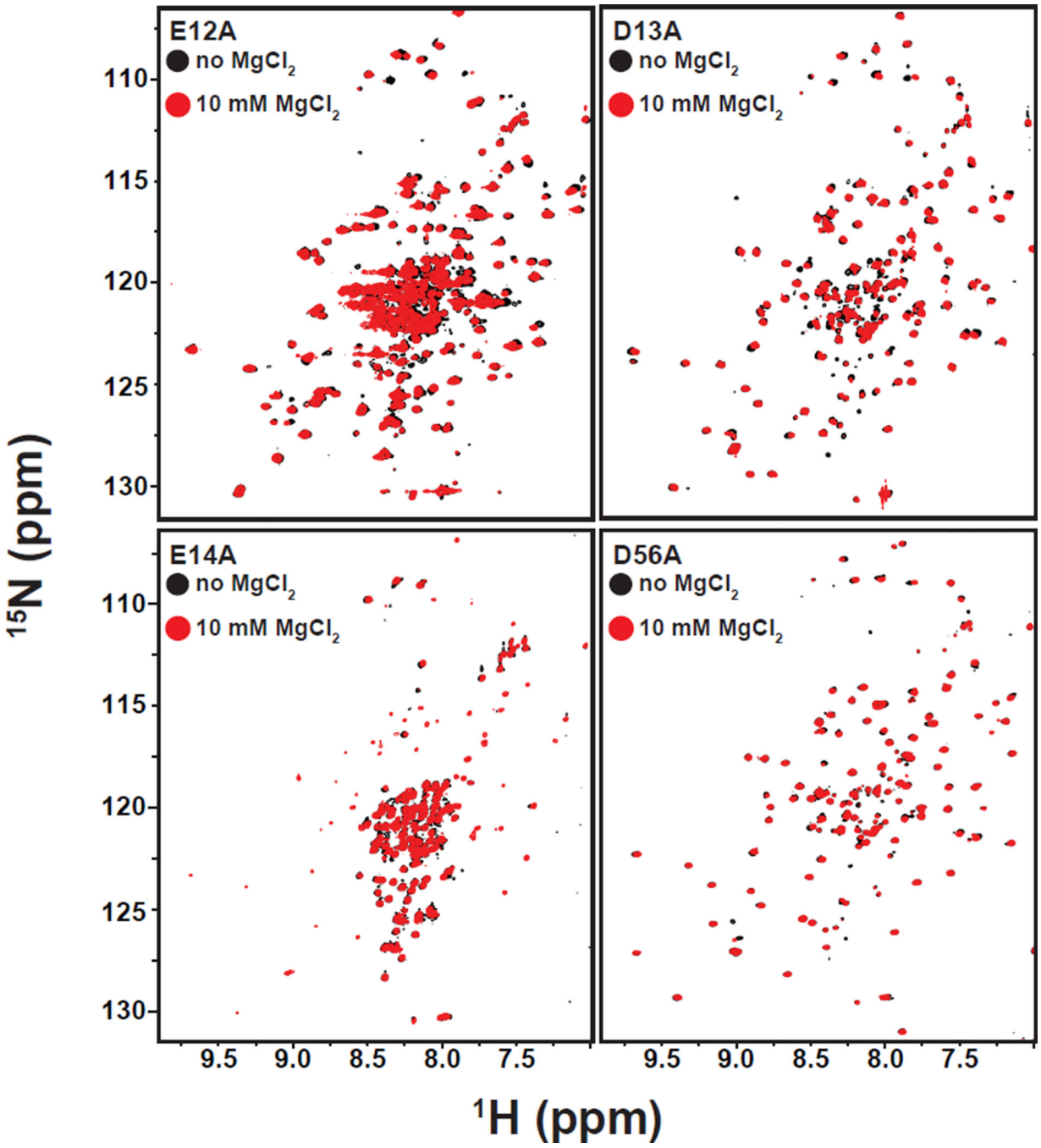

Figure 7. Structural effects of point mutations to the acidic Mg^{2+}-binding residues of EL_LovR

$^{15}N/^{1}H$ HSQC spectra of EL_LovR E12A, D13A, E14A and D56A are shown in the absence of Mg^{2+} (black) or presence of 10 mM $MgCl_2$ (red). For all four cases, the minimal Mg^{2+}-induced chemical shift changes indicate these residues are all required for Mg^{2+} coordination. Additionally, the improved ^{1}H chemical shift dispersion and uniform peak intensities of D13A and D56A (compared to the other mutants or wildtype protein, Figures 4 and S3) indicate that these proteins have become folded without requiring ligand binding.

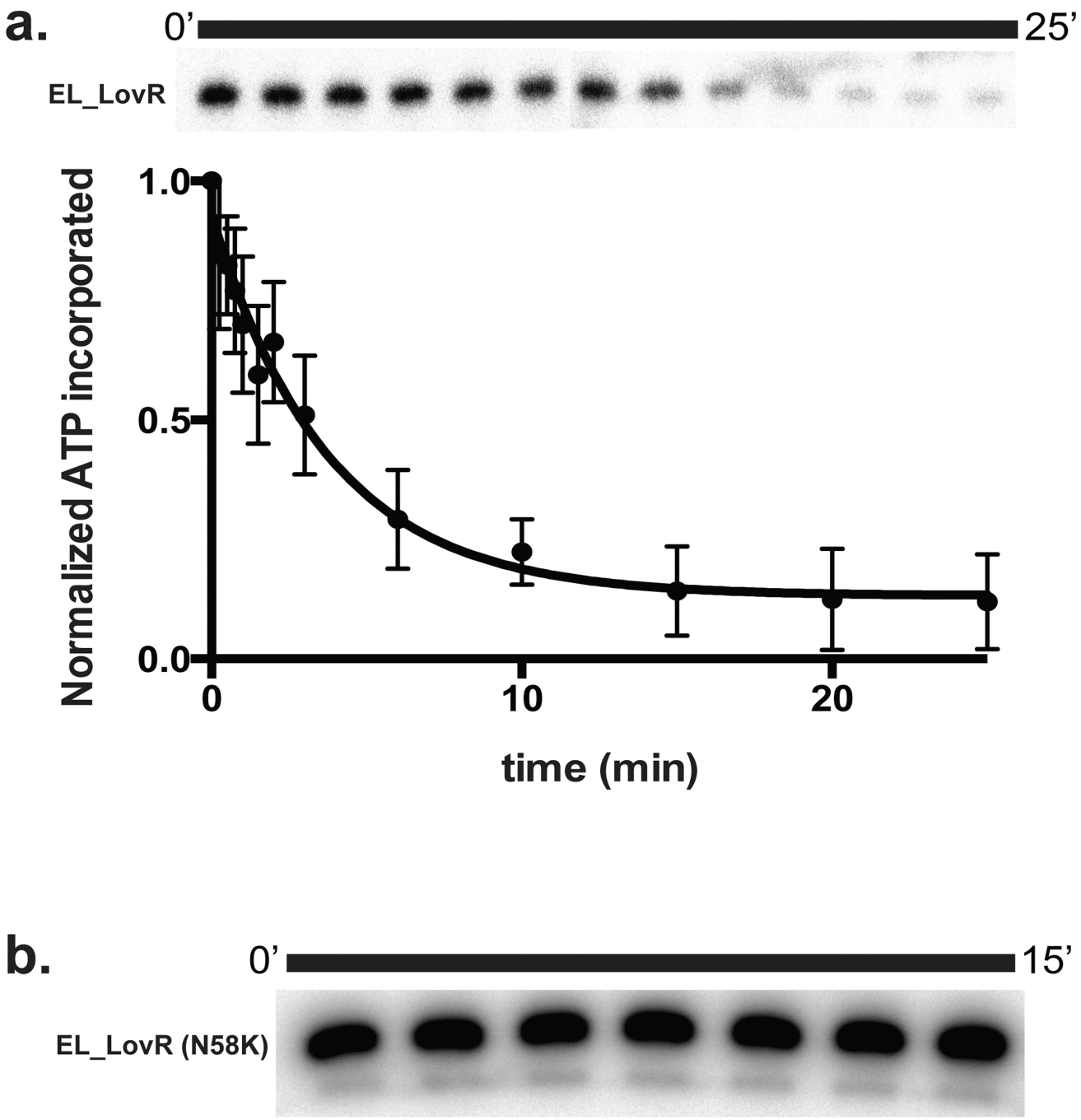

Figure 8. EL_LovR~P is short-lived *in vitro*

a) EL_LovR was phosphorylated by EL368 for 10 min before treatment with 5 mM AMP-PNP to allow the monitoring of phosphate loss over the next 25 min. Normalized levels of incorporated ATP were plotted and fit with a first-order exponential decay to determine the dephosphorylation rate of EL_LovR~P ($k=0.27\ min^{-1}$). Error bars indicate the standard deviation of 3 independent experiments. b) Dephosphorylation assay of EL_LovR (N58K).

Following phosphorylation by the EL368 kinase and addition of AMP-PNP, we observe that EL_LovR (N58K) stably retained the phosphoryl group over 15 min.

Table 1

Statistics for EL_LovR solution structure determination

Structural analysis	
NOE distance restraints	
Unambiguous	1766
Ambiguous	845
Hydrogen bond restraints	50
Dihedral angle restraints	178
Mean rmsd from experimental restraints	
NOE, Å	0.03 ± 0.001
Dihedral angles, °	0.4 ± 0.05
Average no. of	
NOE violations >0.5 Å	0.0
NOE violations >0.3 Å	0.5 ± 0.5
Dihedral violations >5°	0.0
Mean rmsd from idealized covalent geometry	
Bond lengths, Å	0.004 ± 0.0001
Bond angles, °	0.5 ± 0.01
Impropers, °	1.4 ± 0.06
Geometric analysis of residues 8–118	
rmsd from the mean	
Backbone atoms, Å	0.5 ± 0.1
All heavy atoms, Å	1.1 ± 0.1
Ramachandran analysis (PROCHECK)	
Most-favored region, %	91.2
Additionally allowed region, %	8.5
Generously allowed region, %	0.1
Disfavored region, %	0.2