Received: 16 April 2020 | Revised: 20 July 2020 | Accepted: 23 July 2020

DOI: 10.1002/pro.3920

PROTEIN STRUCTURE REPORTS

THE PROTEIN SOCIETY WILEY

Structure and function of the juxtamembrane GAF domain of potassium biosensor KdpD

Shivesh Kumar[1,2] | **Richard E. Gillilan**[3] | **Dinesh A. Yernool**[2]

[1]Department of Biochemistry, Duke University Medical Center, Durham, North Carolina

[2]Department of Biological Sciences, Purdue University, West Lafayette, Indiana

[3]Macromolecular Diffraction Facility, Cornell High Energy Synchrotron Source (MacCHESS), Cornell University, Ithaca, New York

Correspondence
Shivesh Kumar, Department of Biochemistry, Duke University Medical Center, Durham, NC 27710.
Email: shivesh.kumar@duke.edu

Funding information
Empire State Development's Division of Science, Technology and Innovation; NIGMS, Grant/Award Number: 1-P30-GM124166-01A1; National Science Foundation, Grant/Award Number: DMR-1829070

Abstract

KdpD/KdpE two-component signaling system regulates expression of a high affinity potassium transporter responsible for potassium homeostasis. The C-terminal module of KdpD consists of a GAF domain linked to a histidine kinase domain. Whereas certain GAF domains act as regulators by binding cyclic nucleotides, the role of the juxtamembrane GAF domain in KdpD is unknown. We report the high-resolution crystal structure of KdpD GAF domain ($KdpD^G$) consisting of five α-helices, four β-sheets and two large loops. $KdpD^G$ forms a symmetry-related dimer, wherein parallelly arranged monomers contribute to a four-helix bundle at the dimer-interface, SAXS analysis of KdpD C-terminal module reveals an elongated structure that is a dimer in solution. Substitution of conserved residues with various residues that disrupt the dimer interface produce a range of effects on gene expression demonstrating the importance of the interface in inactive to active transitions during signaling. Comparison of ligand binding site of the classic cyclic nucleotide-binding GAF domains to $KdpD^G$ reveals structural differences arising from naturally occurring substitutions in primary sequence of $KdpD^G$ that modifies the canonical NKFDE sequence motif required for cyclic nucleotide binding. Together these results suggest a structural role for $KdpD^G$ in dimerization and transmission of signal to the kinase domain.

KEYWORDS

effector module, four-helix bundle, GAF domain, juxtamembrane adapter, potassium biosensor, sensor histidine kinase

1 | INTRODUCTION

In many membrane bound sensor histidine kinases, the input–output signaling is mediated by a juxtamembrane adaptor domain, known as HAMP (Histidine kinase, Adenylyl cyclases, Methyl-accepting chemotaxis, and Phosphatase proteins) domains.[1] In fewer histidine kinases, PAS (Per-ARNT-Sim)[2,3] and GAF (cGMP-binding PDEs, *Anabaena* Adenylyl cyclases, and *Escherichia coli* FhlA)[1,4,5] adaptor domains replace HAMP domains. GAF domains are one of the largest families of small-molecule-binding

Abbreviations: APS, advanced photon source; DHp, dimerization histidine phosphotransfer; FHB, four-helix bundle; HAMP, histidine kinase, adenylyl cyclase, methyl-binding proteins, and phosphatases; HEPES, 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid; NTA, nitrilotriacetic acid; PDB, protein data bank; PDE's, phosphodiesterase's; SAD, single-wavelength anomalous dispersion; SAXS, small-angle X-ray scattering; SeMet, selenomethionine; TCS, two-component system; TEV, tobacco etching virus; TM, transmembrane.

regulatory domains involved in various signal-transduction systems and in observed in all forms of life.[5–8] To date, GAF domains are present in ~184,337 receptors of bacteria, 13,956 of eukaryotes and 4,583 of archaea.[9] In certain nucleotide phosphodiesterases, regulation of enzymatic activity is dependent on cyclic nucleotides binding to GAF domain.[7,8,10]

At 894 residues in length, KdpD is a relatively large protein with at least six distinct domains. Topologically it is comprised of centrally situated four transmembrane helices with large cytoplasmic N-terminal (M1-W395) and C-terminal (R499-M894) modules.[11,12] The cytosolic C-terminal module of KdpD, which is the effector module is further divided into three domains that includes the juxtamembrane GAF-domain, followed by histidine phosphotransfer (DHp) domain (or HisKA) and histidine kinase (HATPase_c) domain,[13,14] together they are referred to as GAF-DHp-HK module.

KdpD functions as a membrane bound potassium (K^+) biosensor.[15–17] Although few structure–function studies have been attempted, the structure and the functional role of the individual domains are poorly delineated and so are the interactions between the domains in regulating signaling.[18–23] Although a previous report suggested that the oligomeric state of KdpD to be at least a dimer,[23] the elements contributing to oligomerization were not identified. Here we show that C-terminal GAF-DHp-HK module of KdpD is a dimer that forms an elongated structure using SAXS analysis. We determined the crystal structure of the KdpD GAF domain, identified a four-helix bundle at the dimer interface that is involved in input–output transition in response to changes in potassium concentration.

2 | RESULTS

2.1 | Structural characterization

E. coli KdpD GAF domain, namely $KdpD^G$ (L515-T646) (Figure 1) was expressed in a soluble form, purified (Figure S1A) and crystallized. The resulting GAF domain incorporated four residues at the N-terminus (sequence: GAMD) and 13 extra residues at the C-terminus (KLAAAL EHHHHHH) arising from cloning. During modeling, seven residues (EHHHHHH) at the C-terminal region were not traced due to poor electron density. The structure of $KdpD^G$ exhibits a GAF-like fold (Figure 2a,b). It contains five alpha helices, four beta-sheets and two loops. Whereas the N-terminal α-helix 1 (L515-V524) and the α-helix 2 (P528-F543) were arranged in antiparallel fashion, the C-terminal α-helix 5 (P625-T646) was situated parallel to N-terminal α-helix 1. Clustering together, these helices form a three-helix bundle. The β-sheets 1, 3 and 4 that forms a distorted convex surface were sandwiched between the three-helix bundle and the two large loops (P560-D571 and G582-Y597). These two large loops bookend the β-sheets that form a convex surface with the intervening α-helix-3 forming a cap over it. This face of $KdpD^G$, which points away from the 3-helix bundle forms a putative ligand-binding pocket as defined in other GAF domains. Few GAF domains, particularly those associated with phosphodiesterase's have been shown to bind cyclic nucleotides at this site. Symmetry-related $KdpD^G$ molecules form a dimer along the crystallographic axis (Figure 2c).

2.2 | GAF domain dimer interface

The core of the observed dimer fashioned out of a four-helix bundle was formed by α-helixes 1 and 5 from each symmetry related molecule. It is buttressed by interactions between the small β-turn (S604-K607) connecting β-sheets 3 and 4 of one molecule and the α-helix 1 from the symmetry related molecule. The dimer buried a total surface area of 1,237 $Å^2$ (Figure 2c) suggesting that this interface has the potential to be a relevant to quaternary structure of the full-length KdpD molecule.

The GAF domain dimer interface was characterized by multiple hydrophobic, electrostatic and water mediated interactions (Figures 2c and 3). Analysis of the dimer

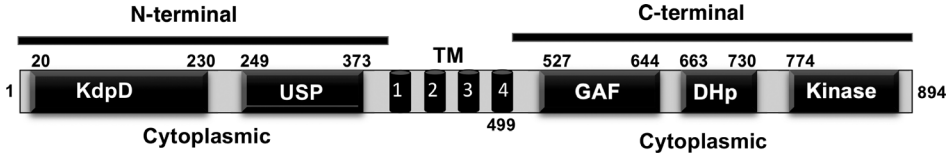

FIGURE 1 Schematic of KdpD delineating various domains. The N-terminal module starts with the namesake KdpD domain containing the classical Walker A and B nucleotide-binding motifs followed by the USP (Universal Stress Protein) domain. The centrally located TM (transmembrane segments 1–4) domain links the N- and C-terminal modules. The C-terminal module consists of the GAF domain, the DHp (dimerization histidine phosphotransfer) domain containing a conserved histidine, and the kinase domain that binds ATP to phosphorylate the conserved histidine in the DHp domain. The numbers above and below the boxes show the start and the end residue of each domain respectively, whereas numbers to the left and right indicate the start and end residue of the full-length KdpD from *E. coli*

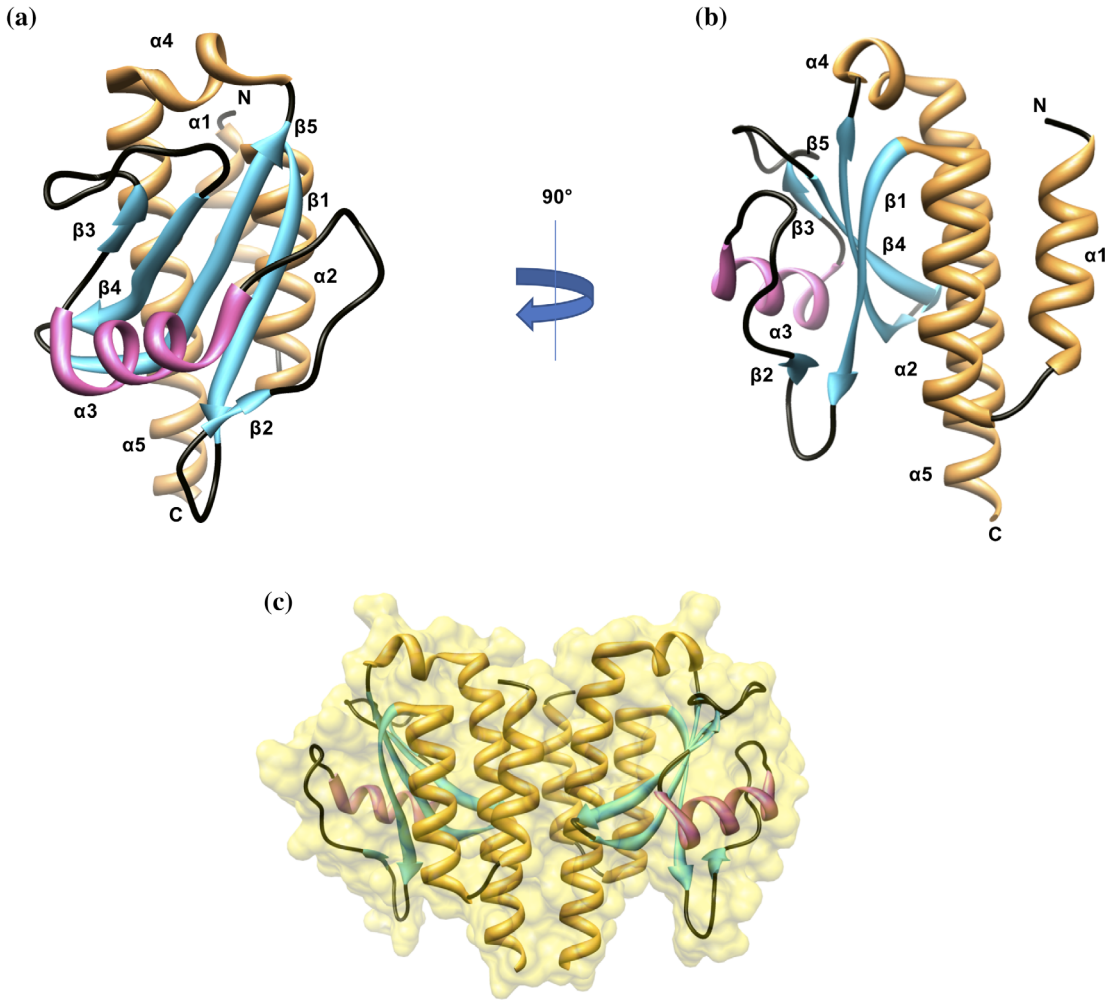

FIGURE 2 Crystal structure of KdpD^{G}. (a) and (b) A ribbon representation of the KdpD GAF domain (KdpD^{G}). Secondary structure elements, loops and termini are labeled and indicated by different colors. Helix α3 (D572-G581), which is located at the putative ligand-binding pocket is colored pink. (c) Representation of a putative dimer formed around the crystallographic axis in space filling and ribbon representation

revealed three layers of interactions stabilizing the interface (Figure 3a). α-helix 5 was a major contributor that participated in all the layers of interactions. Layer 1 involved hydrophobic interactions between the side chains of L515/α-helix 1, T633 and L626/α-helix 5, and an water bridge that connects the hydroxyl group of the side chain of T633 from two molecules of KdpD^{G} (Figure 3b). Layer 2 forms the core of the dimer-interface and involves L637/α-helix 5 from the two protomers (Figure 3c). One-turn further down the helix from the conserved L637/α-helix 5 of layer 2 was the residue N640 which was again one turn above E643 and R644. These residues project their side chains out of the same face of the α-helix 5 into the potential dimer interface. N640, E643 and R644 formed a thicket of interactions between protomers in layer 3 (Figure 3c). E643/α-helix 5 of one protomer interacts with R644/α-helix 5 of the other protomer and vice versa. N640 was involved in water-mediated interaction with R644 of its dimeric counterpart. The residues involved in dimer interface formation (L515, T633, L637, E643 and R644) are highly conserved (Figure 4).

2.3 | Significance of the GAF domain dimer interface

Thus far the data suggests formation of a putative dimer, wherein some residues involved in the dimer interface are conserved among various genera. To evaluate the role of these residues, specific changes were introduced into full-length KdpD, and the variants were assayed for KdpD mediated gene expression using reporter β-galactosidase assays. Wild-type KdpD turns-on and switches-off expression of kdpFABC operon at 0 mM and 10 mM K^{+} concentration, respectively as demonstrated by reporter β-galactosidase assays (Figure 5). Truncation of

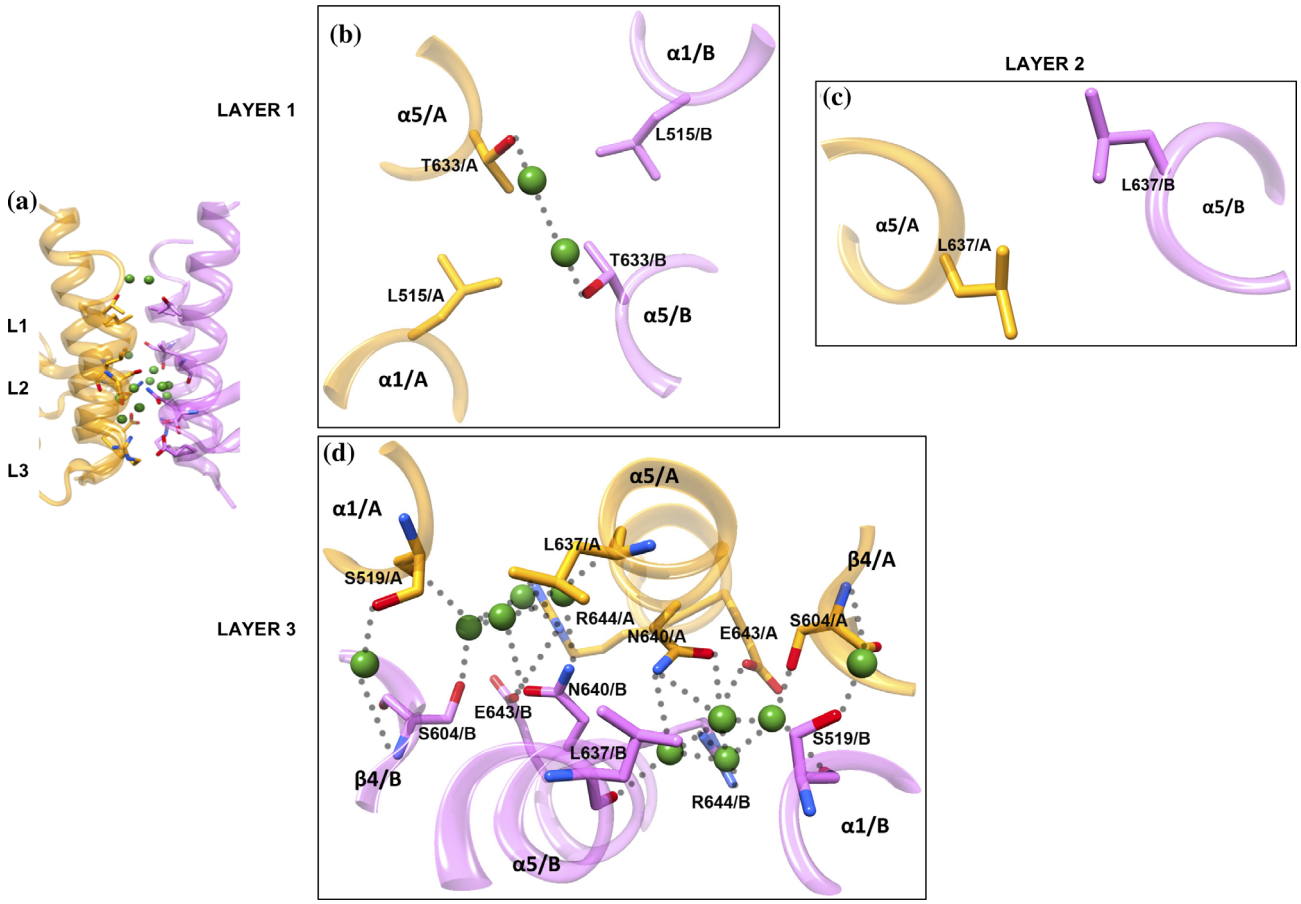

FIGURE 3 Structural details of the central four-helix bundle at the putative dimer interface. (a) The four-helix bundle made of α-helices 1 and 5 from each protomer (colored pink and yellow) is stabilized by hydrophobic, electrostatic, and water-mediated interactions. L1, L2 and L3 represent three layers of interactions at the dimer interface. Residues involved in the dimer interface and the water molecules are shown as sticks and green spheres, respectively. (b) Close view of layer 1 indicates the involvement of L515/ α1 and T633/ α5. The water molecules mediate interactions of T633/ α5 interacts with its symmetry-related counterpart. (c) Layer 2 involves hydrophobic interaction between L637/α5 of two protomers. (d) Layer 3 involves multiple residues involved in an intricate network of electrostatic, hydrophobic and water-mediated interactions

side chain of T633 in layer 1 by introduction of alanine at that position (T633A) has no effect control kdpFABC expression. Introduction of a bulky hydrophobic side chain (T633W) interferes with switching off gene expression at 10 mM K^+ while the induction of expression at 0 mM K^+ remains at wild-type level. Introduction of large charged residues at this position (T633E and T633R) negates both the capability of KdpD to switch-on and switch-off gene expression. Similar pattern of effects on gene expression were observed for substitutions of L637 of layer 2 and the residues in layer 3 namely E643 and R644 (Figure 4).

2.4 | Putative ligand binding pocket

Analysis of $KdpD^G$ sequence shows the lack of conserved NKFDE motif (Figure 6a),[73] common to cyclic nucleotide-binding GAF domains of phosphodiesterases. NKFDE motif was shown to be required for the folding stability of the GAF domain or stabilization of the cyclic nucleotide-binding pocket in cyclases and PDEs.[8,24–27] The structural comparison of $KdpD^G$ with cAMP-bound PDE10A (PDB ID 2ZMF; Figure 6b) and cGMP-bound PDE6C (PDB ID 3DBA; Figure 6c) revealed that the loop connecting β1-α3 (L550-D572) occludes the putative binding pocket in $KdpD^G$.

2.5 | Analysis of SAXS data

KdpD C-terminal full-length construct, the GAF-DHp-HK module, which we termed as $KdpD^{GDK}$ was expressed, purified (Figure S1B) used for SAXS data collection (Table 1). Computed molecular mass and radius of gyration values for $KdpD^{GDK}$ show linear Guinier plots with

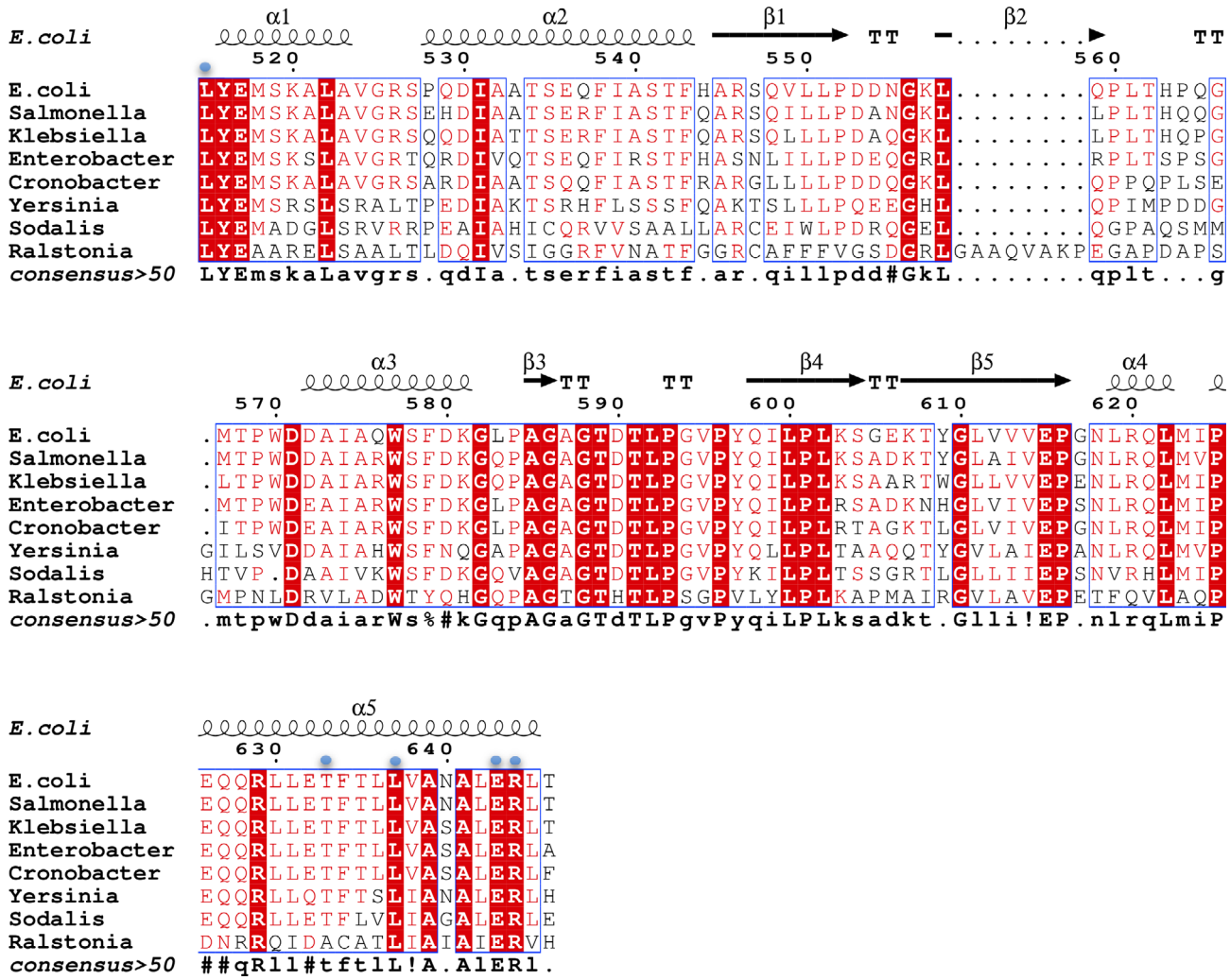

FIGURE 4 The multiple sequence alignment of KdpD GAF domains ($KdpD^G$) from various bacteria. The residues involved in the dimer interface formation are indicated by black dots. The conserved residues are boxed. The α-helices and β-strands shown above the alignment are derived from the structure for *E. coli* $KdpD^G$ (this study), whereas loops and unstructured regions are not shown

expected values and no significant variation across the range of concentrations studied (Figure 7d). The pair distance distribution function, $P(r)$, is slightly bimodal indicating at least two separated domains (Figure 7c). Shape reconstruction produced good statistics (mean value of NSD = 0.664, $\chi^2 = 1.062$ for the reference) with no rejected structures. The resulting shape, however, did not match any biophysically meaningful arrangement of domains. Analysis of the $P(r)$ function using the AMBIMETER program reveals the data can match 309 shape categories producing a poor Ambiguity score of 2.49. Despite the apparent ambiguity, AMBIMETER reported a crude "best" 7-bead structure that matches the expected structural features of $KdpD^{GDK}$ surprisingly well.

Currently, we lack an experimentally determined 3D-structure of $KdpD^{GDK}$. Therefore, we constructed a composite model of $KdpD^{GDK}$ using coordinates from *E. coli* histidine kinase (PDB ID 4CTI) from which the coordinates for HAMP domain were excluded while retaining those of DHP and HK domains and the crystal structure of $KdpD^G$ determined in this study. The sequences of 4CTI and $KdpD^{GDK}$ were aligned to the conserved motif SHDLRTPLT to determine the number of residues between the joining GAF and 4CTI helices (Figure S3). GAF dimer was manually docked into the crude 7-bead structure while attempting to preserve reasonable continuity of helices. At this level of resolution, sequence composition of the *E. coli* DHp Kinase domain is preserved, so no attempt is made to construct a true homology model here, nor to refine connecting helices. The GAF domain is situated in the expected location with a slight asymmetric tilt, but it appears to preserve the natural twist of the helix pair best seen from the top view (Figure 7b). The computed scattering profile from the manually docked model fits the experimental profile well ($\chi^2 = 1.368$).

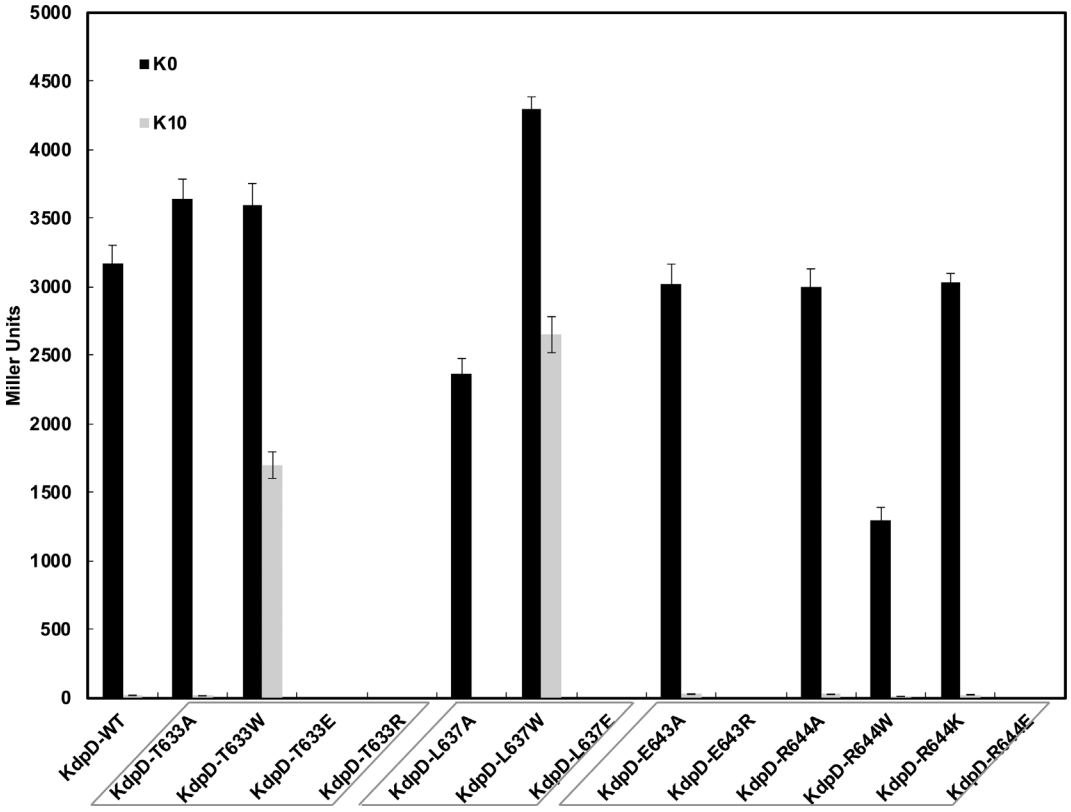

FIGURE 5 Role of residues at the interface of the four-helix bundle of $KdpD^{G}$ in KdpD-KdpE mediated signaling. The amino acid substitutions introduced by mutating full-length KdpD gene used in this analysis are shown below the *X*-axis, the boxes represent Layers 1–3 from left to right. The reporter β-galactosidase production is expressed as Miller units in media without K^+ (K0) and in 10 mM K^+ (K10) conditions permissive and nonpermissive for gene expression respectively. Error bars represent SEM of three independent measurements

3 | DISCUSSION

KdpD/KdpE two-component signal transduction system is activated under conditions of potassium limitation and/or under salt-stress to aid in regaining homeostasis.[13,28] Additionally, it is used as an adaptive regulator to promote the intracellular survival, in a few pathogenic bacteria.[29–32] The effector module of KdpD C-terminal region ($KdpD^{GDK}$) is a GAF-linked histidine kinase, where the juxtamembrane GAF domain ($KdpD^{G}$) connects transmembrane segment to DHp domain. Such a topological organization is equivalent to that of the conventional HAMP domain containing histidine kinases wherein, the HAMP domain acts as a juxtamembrane signal transducer between the ligand binding domain, transmembrane segments, and the DHp-HK domains responsible for histidine autophosphorylation activity. Such histidine kinase activity mostly occurs in trans therefore, they are at least dimers. $KdpD^{G}$ belongs to the large GAF family of proteins, some of its members are capable of binding ligands to regulate enzymatic activity and promote oligomerization of other signaling proteins.[8,25,26,33–38] Therefore, potential exists for dual roles, namely ligand binding and dimerization of the GAF domain which has not been defined for KdpD and other GAF containing histidine kinases. Although $KdpD^{G}$ lacks conserved motif for nucleotide binding and therefore unlikely to bind cyclic nucleotides, other potential ligand(s) and ligand binding dependent regulation remain to be discovered.

To understand the structure–function relationships of $KdpD^{G}$, we have identified, expressed, and purified $KdpD^{G}$ (L515-T646) and $KdpD^{GDK}$ (L515-M894). However multiple attempts to express $KdpD^{DK}$ with varying domain boundaries were futile. Crystallization was successful for Kdp^{G}, but not for $KdpD^{GDK}$ resulting in a high-resolution structure of the former wherein a putative dimer interface mediated by α-helix 1 (L515-V524) and α-helix 5 (P625-T646) forms a four-helix bundle. Although purified $KdpD^{G}$ was not a dimer in solution (data not shown), The SAXS analysis reveals envelope that fits an elongated composite model representing two $KdpD^{GDK}$ molecules wherein two $KdpD^{G}$ domains are juxtaposed. Furthermore, in the crystal structure parallelly arranged protomers, and the secondary structure elements involved at the putative dimer interfaces of GAF domains of $KdpD^{G}$ are comparable to dimer

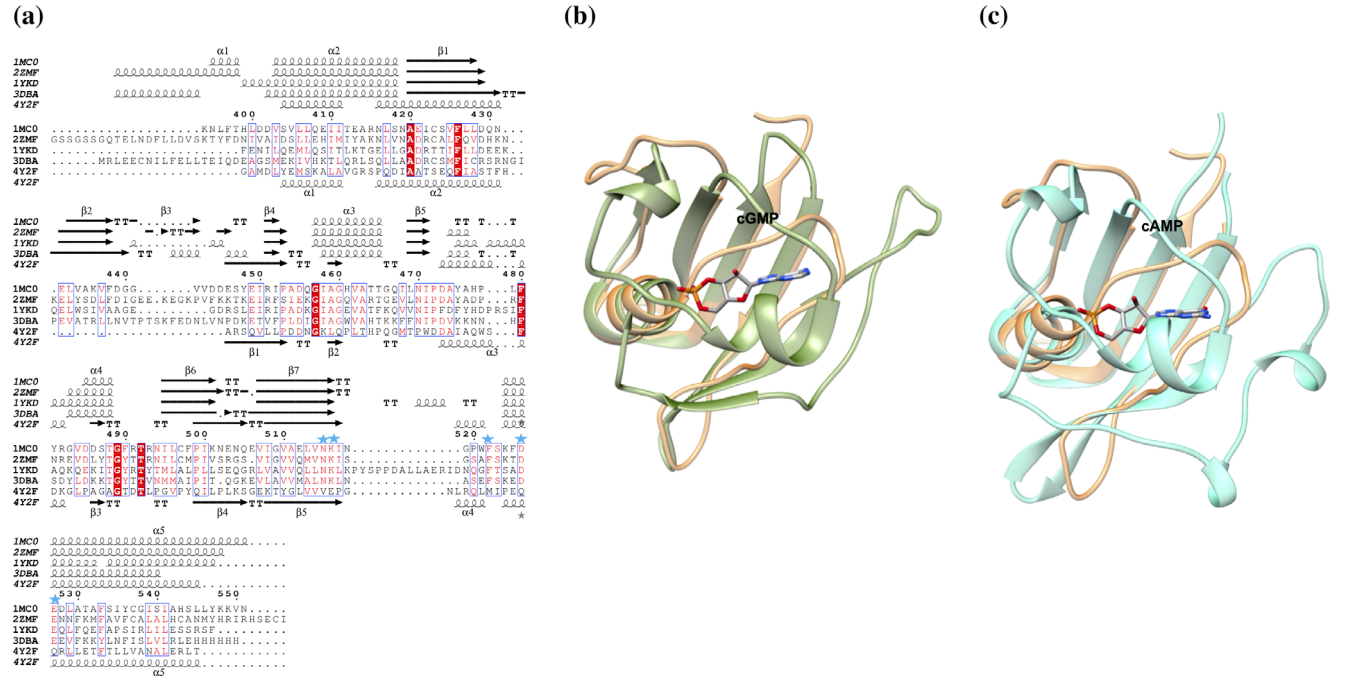

FIGURE 6 Sequence and putative active site analysis. (a) The multiple sequence alignment of cyclic nucleotide binding GAF domains with NKFDE motif marked as blue star. The secondary structure elements are shown for KdpD^{G} (PDB ID 4Y2F) above the alignment. (b). Structural superposition of cGMP-bound GAF-A domain from the photoreceptor phosphodiesterase 6C (PDB ID 3DBA; cartoon in green) on KdpD GAF domain (PDB ID 4Y2F; cartoon in orange). The bound cGMP is shown as a stick model. (c) Structural superposition of cAMP-bound GAF-B domain from the photoreceptor phosphodiesterase 10A (PDB ID 2ZMF; cartoon in cyan) on KdpD GAF domain (PDB ID 4Y2F; cartoon in orange). The bound cAMP is shown as a stick model

TABLE 1 SAXS data

KdpD^{GDK} (mg/ml)	qStart	qEnd	qRg_min	qRg_max	Rg	rsq	I0	MM
0.389	0.01042	0.2833	0.37684	1.27705	36.16584	0.9478	0.18914	76.93601
0.649	0.01274	0.2833	0.46303	1.28384	36.35807	0.95776	0.32628	79.54997
0.811	0.011	0.2833	0.39696	1.29531	36.0913	0.96555	0.38809	75.71918
1.29	0.01042	0.2833	0.38995	1.2998	37.42348	0.98138	0.69296	84.99838
1.62	0.01216	0.2833	0.46219	1.21048	38.0199	0.98394	0.81525	79.62855
1.948	0.01042	0.2833	0.38817	1.29387	37.25261	0.98898	1.04226	84.66034
2.597	0.01042	0.2833	0.39293	1.28792	37.70987	0.99295	1.40474	85.5882
3.24	0.011	0.2833	0.40154	1.28914	36.5082	0.99482	1.64379	80.27699

interfaces reported in GAF containing PDE's. In particular, the arrangement of α-helices 1 and 5 that lead into and out of the GAF-fold of both these multidomain proteins to form a four-helix bundle at the dimer interface. α-helix 1 is a continuum of the transmembrane helix IV in KdpD (Figure S2), while α-helix 1 from PDE2A GAF-B is continuum of a longer α-helix emerging from the N-terminally situated tandem GAF-A domain (Figure S4). Similarly, α-helix 5 of the four-helix bundles of GAF domains of KdpD and PDE2A lead out to their respective enzymatic domains (Figures S2[73] and S4).

The more common and better studied HAMP domain histidine kinases also have a four-helix bundle organization involved in input–output transition during signaling.[7,39–45] Various structure based mechanistic models have been developed for HAMP containing effector modules invoke either helix rotation, helix tilting or bending, helix kinks or lever-like motions of the helix to accomplish conformational changes necessary to move from inactive to active states and vice versa.[42,44–49] Crux to these models is the presence of a central four-helix bundle and the reorientation of these helices relative to each other around a

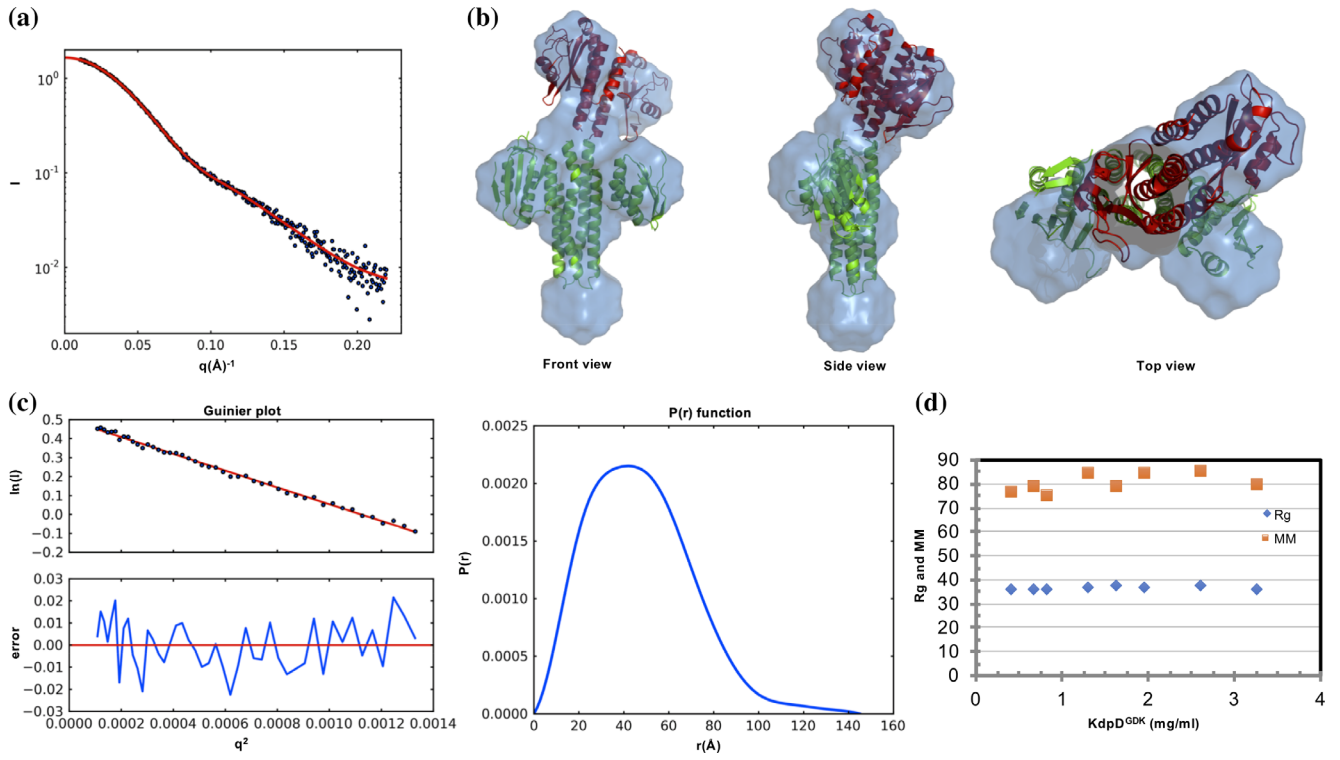

FIGURE 7 Analysis of the C-terminal module ($KdpD^{GDK}$) using SAXS. (a) SAXS scattering curve, (b) SAXS model fitted with models of $KdpD^{G}$ dimer in redand model of the DHp-HK domains from *E. coli* histidine kinase (PDB ID: 4CTI) excluding the HAMP domain at the bottom in green. (c) Guinier plot and Pair-distribution function, *P*(*r*). (d) Concentration dependent changes in Rg and molar mass

shared axis at the dimer interface resulting in altered helix–helix interactions during input–output transitions. Similar rearrangements at the dimer interface between helices constituting the four-helix bundle in GAF containing histidine kinase such as KdpD can be envisioned based on our data including GAF structure, sequence analysis and SAXS analysis-based model showing GAF domains as dimers. If true, the disruption of the observed dimer interface should cause changes in the activity of KdpD, therefore gene expression of the K^+ pump. Functional studies using full-length KdpD shows replacement of residues with alanine that is expected not to disrupt the putative dimer interface maintains wild-type behavior in regulating reporter gene expression. Interestingly, introduction of a bulky side chain residues like tryptophan results in wild-type level of reporter activity at permissive K^+ concentration (0 mM) whereas the ability to completely switch-off gene expression at nonpermissive K^+ concentrations (10 mM) is compromised. In contrast, introduction of charged side chains that is likely to disrupt the interface abrogates gene expression in both permissive and nonpermissive K^+ concentrations indicating inactivation of KdpD mediated signaling. Taken together, the elongated dimeric structural model based on SAXS analysis of $KdpD^{GDK}$, the α-helical structure predicted for sequences preceding and following α-helices 1 and 5 of GAF domain, respectively and the analysis of mutants suggests input–output transitions proposed for HAMP containing histidine kinases are broadly applicable to GAF containing histidine kinases.

4 | MATERIALS AND METHODS

4.1 | Cloning, overexpression, and purification of KdpD-GAF domain ($KdpD^{G}$) and KdpD C-terminal domain containing GAF-DHp-HK module ($KdpD^{GDK}$)

DNA isolated from E. coli str. K-12 substr. MG1655 was used as a source of KdpD gene (accession # AAC73788). Sequences encoding $KdpD^{G}$ (L515-T646) and $KdpD^{GDK}$ (L515-M894) were cloned into the Nco1 and Xho1 sites of pHIS-parallel1 vector,[50] which introduces a His6-TeV protease site tag at the N-terminal of the translated fusion protein (Tables S1 and S2). The entirety of KdpD (KdpD-FL) gene was amplified by PCR and cloned into pTEVGH11 vector.[51] This construct can express a fusion protein wherein a GFP-His11 tag is fused to the C-terminus KdpD-FL protein. Site directed mutagenesis was carried out on this construct capable of expressing KdpD-GFP-His11 using overlap PCR method to introduce single residue changes for structure–function studies.[52]

For protein production, *E. coli* C43 cells containing $KdpD^{G}$ and $KdpD^{GDK}$ constructs grown in LB media to OD_{600} of 0.7 were induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at 37°C and harvested after 5 hr. For labeling $KdpD^{G}$ with Se-Met the protocol described by Hendrickson et al. was followed.[53] Pellets of $KdpD^{G}$ expressing cells resuspended in buffer A (10 mM HEPES pH 7.5, 0.15 M NaCl) was supplemented with 0.1 mg DNase and protease inhibitor cocktail (Sigma Aldrich Co.). The cells were lysed using homogenizer (Avestin Corp.) at 15,000 psi, clarified by centrifugation to obtain clarified lysate, which was passed over Ni-NTA column to obtain purified His_6-TeV-$KdpD^{G}$. The purified His_6-TeV-$KdpD^{G}$ was incubated with 1:10 wt/wt of TEV protease overnight at room temperature and the mixture was precipitated by adding ammonium sulfate to 75% saturation. The precipitate was dissolved in buffer B (10 mM HEPES pH 7.5, 50 mM NaCl) and subjected to size-exclusion chromatography using HiLoad 16/60 S200 column (GE Healthcare) equilibrated with the same buffer. The purified $KdpD^{G}$ was concentrated using Amicon Ultra Centricon tubes (Millipore, Massachusetts). Generally, purified proteins were quantified using Nanodrop 2000c.

$KdpD^{GDK}$ was purified in buffer C (25 mM TRIS pH 8.0 and 0.5 M NaCl, 5 mM β-ME) using Ni-NTA affinity chromatography as described for $KdpD^{G}$. Purified $KdpD^{GDK}$ was subjected to ion-exchange chromatography using Resource Q column (GE Healthcare) in buffer D (25 mM TRIS pH 8.0, 50 mM NaCl, 5 mM β-ME) with linear salt gradient (0.05–1 M NaCl). The fractions containing highest amounts of protein was concentrated and further purified by size-exclusion chromatography using HiLoad 16/60 S200 column (GE Healthcare) in buffer E (25 mM HEPES pH 7.5, 150 mM NaCl, 2 mM β-ME). The peak fractions from size-exclusion chromatography were used for SAXS analysis.

4.2 | Crystallization, data collection, and structure solution of GAF domain

Crystallization of $KdpD^{G}$ was performed using hanging-drop method at 20°C by mixing protein (15 mg/ml) with well solution (30–35% PEG 3350, 0.1 M Bis-TRIS pH 6.0–6.5) at a 1:1 ratio. The crystals were cryoprotected by soaking in 35% PEG 3350, 0.1 M Bis-TRIS pH 6.5, 30% glycerol before flash-freezing in liquid nitrogen. Diffraction data from native crystal was collected at beamline 21-IDD, Advanced Photon Source and processed using iMOSFLM.[54] Structure determination by molecular replacement method failed due to very low sequence conservation to other GAF domains. Therefore, experimental SAD phasing was conducted using crystal of Se-Met labeled $KdpD^{G}$, purified and crystallized using methods outlined for native protein. The full 360° diffraction data set (d_{limit} = 1.54 Å) was collected at selenium peak-wavelength (0.97932 Å) at beamline 21-IDD, Advanced Photon Source. The data collection and refinement statistics are shown in Table 1. The crystal belonging to space group P21212 contains one molecule in asymmetric unit with solvent content and Matthews's coefficient of 34.52% and 1.88 $Å^3$/Da, respectively. The overall anomalous signal-to-noise ratio is 3.44. Substructure solution, phasing and density modification followed by partial model building were performed using *PHENIX AutoSol* and *AutoBuild* wizard.[55] A total of four selenium sites were located including one site arising from the start codon derived from the vector. The Bayesian estimate of map CC and figure of merit (FOMs) were 0.52 and 0.58, respectively. Iterative model building and refinement were performed in COOT[56] and PHENIX.[55] The final model has R-factor and R-free of 18.49% and 20.56%, respectively (Table 2) which were within the range of average values for structures refined at these resolutions.[57] The final model has 98.57% of the residues in Ramachandran favored region and 1.43% of the residues in allowed region. The structures were validated using comprehensive validation tools in PHENIX.[58,59] Graphics were generated using PyMOL (http://www.pymol.org) and Chimera.[60]

4.3 | SAXS data collection, processing, and modeling

$KdpD^{GDK}$ at concentrations of 0.389, 0.649, 0.811, 1.29, 1.62, 1.948, 2.597, 3.24 mg/ml were prepared in buffer E and centrifuged at 14,000 rpm for 15 min prior to data collection. SAXS data were collected on CHESS beamline F2 at 9.881 keV (1.2563 Å, the tantalum edge) at 1×10^{10} photons/s. The X-ray beam was collimated to 250×250 μm^2 and was centered on a 2 mm diameter quartz capillary with 10 μm thick walls (Hampton Research, Aliso Viejo, CA). To eliminate air scatter from the capillary tube, full X-ray flight path and beam stop were maintained in vacuum. Using a Hudson SOLO single-channel pipetting robot (Hudson Robotics Inc., Springfield, NJ), 20 μl samples were delivered from a 96-well plate to the capillary. The sample plugs were oscillated in the X-ray beam using a computer-controlled syringe pump (Aurora Biomed, Vancouver, Canada) to reduce radiation damage. Images were collected on a Pilatus 100 K-S detector (Dectris, Baden, Switzerland). Sample and buffer solutions were normalized to equivalent exposure before subtraction using beamstop photodiode counts. The complete automated Bio-SAXS system is described elsewhere.[61]

TABLE 2 Data collection and refinement statistics

Data sets	KdpD^G (Se)	KdpD^G (native)
Crystallographic data		
X-ray source	21-IDD, APS, Argonne, IL	Rigaku MSC, Purdue University
Wavelength (Å)	0.97932	1.54178
Space group	P21212	P21212
Unit cell parameters (Å)		
a, *b*, *c*	35.31, 57.61, 61.00	32.25, 57.54, 61.06
Resolution range	30.5–1.548 (1.603–1.548)	28.77–1.396 (1.446–1.396)
Completeness (%)	99.97 (100)	98.3 (97.1)
Total no. of reflections	270,303	350,555
No. of unique reflections	18,769	24,876
Wilson B-factor (Å^2)	16.63	17.90
Multiplicity	14.4 (14.1)	14.1 (13.9)
R-sym (%)	0.078 (0.44)	0.073 (1.96)
Mean I/σ (I)	18.75 (6.18)	18.0 (1.7)
Crystal mosaicity (°)	0.23	0.74
Refinement statistics		
R factor (%)	18.49	19.09
Free_R_factor (%)	20.56	21.12
Mean B-factor (Å^2)	22.50	28.40
No. of atoms		
Protein/H_2O	1080/99	1087/141
RMS deviations		
RMS (bonds)	0.008	0.005
RMS (angles)	1.21	1.01
Ramachandran favored (%)	99	99
Ramachandran outliers (%)	0	0

Note: Values in parentheses are for the highest resolution shell. Rfree was calculated with 10% of the reflections selected randomly.

The sample-to-detector distance was calibrated using silver behenate powder (The Gem Dugout, State College, PA). CCD images were reduced to profiles and buffer-subtracted using the BioXTAS RAW software.[62] The q-space range used, where $q = 4\pi\sin(\theta)/\lambda$, with 2θ being the scattering angle, was determined using the Guinier plot as a guide. Generally, $q_{min} = 0.01$ Å^{-1} and $q_{max} = 0.27$ Å^{-1}. Radius of gyration (R_g) was calculated using both the method of Guinier[63] and the inverse Fourier transform (IFT) method as implemented in the GNOM program.[64] In the case of the Guinier method, linear fitting was performed on data having a range of $qR_g < 1.3$ unless otherwise noted. Interactive fitting was performed using the BioXTAS RAW program.[62]

Distance distribution functions were calculated using the program GNOM over a range of D_{max} values, examining the total estimate scores and quality of fit at each step.[65] Values of D_{max} chosen for comparison of $P(r)$ functions were the smallest values that yielded smooth decay of the curve to zero while staying below the Shannon limit set by the experimental apparatus: $D_{max} < \pi/q_{min}$. Ambiguity assessment of the final scattering profile was performed using the AMBIMETER program.[63] Shape reconstruction was performed by running 10 independent DAMMIF calculations[66] and building a consensus envelope using the DAMAVER program.[67] Manual docking of model domains into envelopes was performed with PyMOL (http://www.pymol.org). Scattering profiles for models were generated and compared with experimental data using the CRYSOL program.[68]

4.4 | β-Galactosidase activity assay

Assays for β-galactosidase were performed in strain HAK006,[69] which possess a*kdpFABC* promoter-*lacZ* fusion gene on the chromosome. Cells transformed with plasmids expressing the wild-type KdpD or its GAF domain dimer interface variants (Table S1) were grown in minimal medium[70] containing KCl at various concentrations (0.0 and 10 mM) and assays were performed as described previously.[71] All assays were done in triplicate.

ACKNOWLEDGMENTS

We are grateful to Zdzislaw Wawrzak at 21-IDD, LS-CAT beamline, APS for help with data collection and Prof. Pei Zhou for advice on the manuscript. The SAXS data collection and analysis has been conducted at the Center for High Energy X-ray Sciences (CHESX), which is supported by the National Science Foundation under award DMR-1829070, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by NIGMS award 1-P30-GM124166-01A1 and NYSTAR.

AUTHOR CONTRIBUTIONS

Shivesh Kumar: Conceptualization; data curation; formal analysis; methodology; validation; visualization; writing-original draft; writing-review and editing. **Richard E. Gillilan:** Methodology; resources; software; validation; visualization; writing-review and editing. **Dinesh A. Yernool:** Conceptualization; data curation; formal

analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing-review and editing.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ACCESSION CODES

Structure factors and atomic coordinates have been deposited in the Protein Data Bank under accession codes 4QPR and 4Y2F for SeMet labeled and native $KdpD^G$ structures, respectively.

ORCID

Shivesh Kumar https://orcid.org/0000-0002-3414-4091

REFERENCES

1. Aravind L, Ponting CP. The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. FEMS Microbiol Lett. 1999;176:111–116.
2. Möglich A, Ayers RA, Moffat K. Structure and signaling mechanism of Per-ARNT-Sim domains. Structure. 2009;17:1282–1294.
3. Henry JT, Crosson S. Ligand-binding PAS domains in a genomic, cellular, and structural context. Annu Rev Microbiol. 2011;65:261–286.
4. Martinez SE, Beavo JA, Hol WGJ. GAF domains: Two-billion-year-old molecular switches that bind cyclic nucleotides. Mol Interv. 2002;2:317–323.
5. Zoraghi R, Corbin JD, Francis SH. Properties and functions of GAF domains in cyclic nucleotide phosphodiesterases and other proteins. Mol Pharmacol. 2004;65:267–278.
6. Cann MJ. Sodium regulation of GAF domain function. Biochem Soc Trans. 2007;35:1032–1034.
7. Dunin-Horkawicz S, Lupas AN. Comprehensive analysis of HAMP domains: Implications for transmembrane signal transduction. J Mol Biol. 2010;397:1156–1174.
8. Martinez SE, Wu AY, Glavas NA, et al. The two GAF domains in phosphodiesterase 2A have distinct roles in dimerization and in cGMP binding. Proc Natl Acad Sci U S A. 2002;99:13260–13265.
9. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: Identification of signaling domains. Proc Natl Acad Sci U S A. 1998;95:5857–5864.
10. Francis SH, Bessay EP, Kotera J, et al. Phosphorylation of isolated human phosphodiesterase-5 regulatory domain induces an apparent conformational change and increases cGMP binding affinity. J Biol Chem. 2002;277:47581–47587.
11. Parkinson JS, Kofoid EC. Communication modules in bacterial signaling proteins. Annu Rev Genet. 1992;26:71–112.
12. Zimmann P, Puppe W, Altendorf K. Membrane topology analysis of the sensor kinase KdpD of *Escherichia coli*. J Biol Chem. 1995;270:28282–28288.
13. Heermann R, Jung K. The complexity of the 'simple' two-component system KdpD/KdpE in *Escherichia coli*. FEMS Microbiol Lett. 2010;304:97–106.
14. El-Gebali S, Mistry J, Bateman A, et al. The Pfam protein families database in 2019. Nucleic Acids Res. 2019;47:D427–D432.
15. Walderhaug MO, Polarek JW, Voelkner P, et al. KdpD and KdpE, proteins that control expression of the kdpABC operon, are members of the two-component sensor-effector class of regulators. J Bacteriol. 1992;174:2152–2159.
16. Epstein W. Kdp, a bacterial P-type ATPase whose expression and activity are regulated by turgor pressure. Acta Physiol Scand Suppl. 1992;607:193–199.
17. Heermann R, Altendorf K, Jung K. The N-terminal input domain of the sensor kinase KdpD of *Escherichia coli* stabilizes the interaction between the cognate response regulator KdpE and the corresponding DNA-binding site. J Biol Chem. 2003;278:51277–51284.
18. Rothenbucher MC, Facey SJ, Kiefer D, Kossmann M, Kuhn A. The cytoplasmic C-terminal domain of the *Escherichia coli* KdpD protein functions as a K^+ sensor. J Bacteriol. 2006;188:9.
19. Zimmann P, Steinbrügge A, Schniederberend M, Jung K, Altendorf K. The extension of the fourth transmembrane helix of the sensor kinase KdpD of *Escherichia coli* is involved in sensing. J Bacteriol. 2007;189:7326–7334.
20. Jung K, Altendorf K. Truncation of amino acids 12–128 causes deregulation of the phosphatase activity of the sensor kinase KdpD of *Escherichia coli*. J Biol Chem. 1998;273:17406–17410.
21. Heermann R, Lippert M-L, Jung K. Domain swapping reveals that the N-terminal domain of the sensor kinase KdpD in *Escherichia coli* is important for signaling. BMC Microbiol. 2009;9:133.
22. Schramke H, Tostevin F, Heermann R, Gerland U, Jung K. A dual-sensing receptor confers robust cellular homeostasis. Cell Rep. 2016;16:213–221.
23. Heermann R, Altendorf K, Jung K. The turgor sensor KdpD of *Escherichia coli* is a homodimer. Biochim Biophys Acta. 1998;1415:114–124.
24. Bruder S, Linder JU, Martinez SE, Zheng N, Beavo JA, Schultz JE. The cyanobacterial tandem GAF domains from the cyaB2 adenylyl cyclase signal via both cAMP-binding sites. Proc Natl Acad Sci U S A. 2005;102:3088–3092.
25. Gross-Langenhoff M, Hofbauer K, Weber J, Schultz A, Schultz JE. cAMP is a ligand for the tandem GAF domain of human phosphodiesterase 10 and cGMP for the tandem GAF domain of phosphodiesterase 11. J Biol Chem. 2006;281:2841–2846.
26. Heikaus CC, Pandit J, Klevit RE. Cyclic nucleotide binding GAF domains from phosphodiesterases—Structural and mechanistic insights. Structure. 2009;17:1551–1557.
27. Turko IV, Haik TL, McAllister-Lucas LM, Burns F, Francis SH, Corbin JD. Identification of key amino acids in a conserved cGMP-binding site of cGMP-binding phosphodiesterases. A putative NKXnD motif for cGMP binding. J Biol Chem. 1996;271:22240–22244.
28. Alegado R, Chin C-Y, Monack D, Tan M-W. The two-component sensor kinase KdpD is required for *Salmonella typhimurium* colonization of *Caenorhabditis elegans* and survival in macrophages. Cell Microbiol. 2011;13:1618–1637.
29. Freeman ZN, Dorus S, Waterfield NR. The KdpD/KdpE two-component system: Integrating K+ homeostasis and virulence. PLoS Pathog. 2013;9:e1003201.
30. Feinbaum RL, Urbach JM, Liberati NT, et al. Genome-wide identification of *Pseudomonas aeruginosa* virulence-related genes using a *Caenorhabditis elegans* infection model. PLoS Pathog. 2012;8:e1002813.

31. Parish T, Smith DA, Kendall S, Casali N, Bancroft GJ, Stoker NG. Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. Infect Immun. 2003;71:1134–1140.
32. Xue T, You Y, Hong D, Sun H, Sun B. The *Staphylococcus aureus* KdpDE two-component system couples extracellular K+ sensing and Agr signaling to infection programming. Infect Immun. 2011;79:2154–2167.
33. Anantharaman V, Koonin EV, Aravind L. Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. J Mol Biol. 2001;307: 1271–1292.
34. Aravind L, Ponting CP. The GAF domain: An evolutionary link between diverse phototransducing proteins. Trends Biochem Sci. 1997;22:458–459.
35. Zoraghi R, Bessay EP, Corbin JD, Francis SH. Structural and functional features in human PDE5A1 regulatory domain that provide for allosteric cGMP binding, dimerization, and regulation. J Biol Chem. 2005;280:12051–12063.
36. Muradov KG, Boyd KK, Martinez SE, Beavo JA, Artemyev NO. The GAFa domains of rod cGMP-phosphodiesterase 6 determine the selectivity of the enzyme dimerization. J Biol Chem. 2003;278:10594–10601.
37. Okada D, Asakawa S. Allosteric activation of cGMP-specific, cGMP-binding phosphodiesterase (PDE5) by cGMP. Biochemistry. 2002;41:9672–9679.
38. Rybalkin SD, Rybalkina IG, Shimizu-Albergine M, Tang X-B, Beavo JA. PDE5 is converted to an activated state upon cGMP binding to the GAF A domain. EMBO J. 2003;22:469–478.
39. Parkinson JS. Signaling mechanisms of HAMP domains in chemoreceptors and sensor kinases. Annu Rev Microbiol. 2010;64: 101–122.
40. Flack CE, Parkinson JS. A zipped-helix cap potentiates HAMP domain control of chemoreceptor signaling. Proc Natl Acad Sci U S A. 2018;115:E3519–E3528.
41. Butler SL, Falke JJ. Cysteine and disulfide scanning reveals two amphiphilic helices in the linker region of the aspartate chemoreceptor. Biochemistry. 1998;37:10746–10756.
42. Hulko M, Berndt F, Gruber M, et al. The HAMP domain structure implies helix rotation in transmembrane signaling. Cell. 2006;126:929–940.
43. Ames P, Zhou Q, Parkinson JS. Mutational analysis of the connector segment in the HAMP domain of Tsr, the *Escherichia coli* serine chemoreceptor. J Bacteriol. 2008;190:6676–6685.
44. Zhou Q, Ames P, Parkinson JS. Mutational analyses of HAMP helices suggest a dynamic bundle model of input–output signalling in chemoreceptors. Mol Microbiol. 2009;73:801–814.
45. Swain KE, Falke JJ. Structure of the conserved HAMP domain in an intact, membrane-bound chemoreceptor: A disulfide mapping study. Biochemistry. 2007;46:13684–13695.
46. Cheung J, Hendrickson WA. Structural analysis of ligand stimulation of the histidine kinase NarX. Structure. 2009;17:190–201.
47. Falke JJ, Hazelbauer GL. Transmembrane signaling in bacterial chemoreceptors. Trends Biochem Sci. 2001;26:257–265.
48. Moukhametzianov R, Klare JP, Efremov R, et al. Development of the signal in sensory rhodopsin and its transfer to the cognate transducer. Nature. 2006;440:115–119.
49. Khursigara CM, Wu X, Zhang P, Lefman J, Subramaniam S. Role of HAMP domains in chemotaxis signaling by bacterial chemoreceptors. Proc Natl Acad Sci U S A. 2008;105: 16555–16560.
50. 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.
51. Narayanan A, Ridilla M, Yernool DA. Restrained expression, a method to overproduce toxic membrane proteins by exploiting operator–repressor interactions. Protein Sci. 2011;20: 51–61.
52. Urban A, Neukirchen S, Jaeger K-E. A rapid and efficient method for site-directed mutagenesis using one-step overlap extension PCR. Nucleic Acids Res. 1997;25:2227–2228.
53. Hendrickson WA, Horton JR, LeMaster DM. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): A vehicle for direct determination of three-dimensional structure. EMBO J. 1990;9:1665–1672.
54. Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. 2011;D67:271–281.
55. Adams PD, Afonine PV, Bunkóczi G, et al. PHENIX: A comprehensive python-based system for macromolecular structure solution. Acta Crystallogr. 2010;D66:213–221.
56. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr. 2010;D66:486–501.
57. Kleywegt GJ, Brünger AT. Checking your imagination: Applications of the free R value. Structure. 1996;4:897–904.
58. Williams CJ, Headd JJ, Moriarty NW, et al. MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. 2018;27:293–315.
59. Urzhumtseva L, Afonine PV, Adams PD, Urzhumtsev A. Crystallographic model quality at a glance. Acta Crystallogr. 2009; D65:297–300.
60. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—A visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–1612.
61. Nielsen SS, Møller M, Gillilan RE. High-throughput biological small-angle X-ray scattering with a robotically loaded capillary cell. J Appl Cryst. 2012;45:213–223.
62. Hopkins JB, Gillilan RE, Skou S. *BioXTAS RAW*: Improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J Appl Cryst. 2017;50:1545–1553.
63. Petoukhov MV, Svergun DI. Ambiguity assessment of small-angle scattering curves from monodisperse systems. Acta Crystallogr. 2015;D71:1051–1058.
64. Semenyuk AV, Svergun DI. GNOM—A program package for small-angle scattering data processing. J Appl Cryst. 1991;24: 537–540.
65. Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Cryst. 1992;25:495–503.
66. Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J Appl Cryst. 2009;42:342–346.
67. Volkov VV, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. J Appl Cryst. 2003;36:860–864.
68. Svergun D, Barberato C, Koch MHJ. *CRYSOL*—A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Cryst. 1995;28:768–773.

69. Nakashima K, Sugiura A, Mizuno T. Functional reconstitution of the putative *Escherichia coli* osmosensor, KdpD, into liposomes. J Biochem. 1993;114:615–621.
70. Epstein W, Davies M. Potassium-dependent mutants of *Escherichia coli* K-12. J Bacteriol. 1970;101:836–843.
71. Narayanan A, Kumar S, Evrard AN, Paul LN, Yernool DA. An asymmetric heterodomain interface stabilizes a response regulator–DNA complex. Nat Commun. 2014;5:1–9.
72. Buchan DWA, Jones DT. The PSIPRED protein analysis workbench: 20 years on. Nucleic Acids Res. 2019;47:W402–W407.
73. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014;42:W320–W324.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Kumar S, Gillilan RE, Yernool DA. Structure and function of the juxtamembrane GAF domain of potassium biosensor KdpD. *Protein Science*. 2020;29:2009–2021. https://doi.org/10.1002/pro.3920