AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Spectrum

RESEARCH ARTICLE

Phylogenetic Analysis with Prediction of Cofactor or Ligand Binding for *Pseudomonas aeruginosa* PAS and Cache Domains

Andrew Hutchin,[a,b,c,d*] Charlotte Cordery,[a,b,c,e] Martin A. Walsh,[b,c] Jeremy S. Webb,[a,e] Ivo Tews[a,e]

[a]Biological Sciences, Institute for Life Sciences, University of Southampton, Southampton, United Kingdom
[b]Diamond Light Source, Harwell Science and Innovation Campus, Didcot, United Kingdom
[c]Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot, United Kingdom
[d]Structure and Function of Biological Membranes Lab, Université Libre de Bruxelles, Brussels, Belgium
[e]National Biofilms Innovation Centre, University of Southampton, Southampton, United Kingdom

ABSTRACT PAS domains are omnipresent building blocks of multidomain proteins in all domains of life. Bacteria possess a variety of PAS domains in intracellular proteins and the related Cache domains in periplasmic or extracellular proteins. PAS and Cache domains are predominant in sensory systems, often carry cofactors or bind ligands, and serve as dimerization domains in protein association. To aid our understanding of the wide distribution of these domains, we analyzed the proteome of the opportunistic human pathogen *Pseudomonas aeruginosa* PAO1 *in silico*. The ability of this bacterium to survive under different environmental conditions, to switch between planktonic and sessile/biofilm lifestyle, or to evade stresses, notably involves c-di-GMP regulatory proteins or depends on sensory pathways involving multidomain proteins that possess PAS or Cache domains. Maximum likelihood phylogeny was used to group PAS and Cache domains on the basis of amino acid sequence. Conservation of cofactor- or ligand-coordinating amino acids aided by structure-based comparison was used to inform function. The resulting classification presented here includes PAS domains that are candidate binders of carboxylic acids, amino acids, fatty acids, flavin adenine dinucleotide (FAD), 4-hydroxycinnamic acid, and heme. These predictions are put in context to previously described phenotypic data, often generated from deletion mutants. The analysis predicts novel functions for sensory proteins and sheds light on functional diversification in a large set of proteins with similar architecture.

IMPORTANCE To adjust to a variety of life conditions, bacteria typically use multidomain proteins, where the modular structure allows functional differentiation. Proteins responding to environmental cues and regulating physiological responses are found in chemotaxis pathways that respond to a wide range of stimuli to affect movement. Environmental cues also regulate intracellular levels of cyclic-di-GMP, a universal bacterial secondary messenger that is a key determinant of bacterial lifestyle and virulence. We study *Pseudomonas aeruginosa*, an organism known to colonize a broad range of environments that can switch lifestyle between the sessile biofilm and the planktonic swimming form. We have investigated the PAS and Cache domains, of which we identified 101 in 70 *Pseudomonas aeruginosa* PAO1 proteins, and have grouped these by phylogeny with domains of known structure. The resulting data set integrates sequence analysis and structure prediction to infer ligand or cofactor binding. With this data set, functional predictions for PAS and Cache domain-containing proteins are made.

KEYWORDS Cache domain, PAS domain, phylogeny, *Pseudomonas*, cofactors, phylogenetic analysis, sensory transduction processes

Editor Gaurav Sharma, Institute of Bioinformatics and Applied Biotechnology

Address correspondence to Ivo Tews, ivo.tews@soton.ac.uk.

*Present address: Andrew Hutchin, Evotec (UK) Ltd., Abingdon, United Kingdom.

The authors declare no conflict of interest.

Received 26 July 2021
Accepted 21 November 2021
Published 22 December 2021

The Gram-negative bacterium *Pseudomonas aeruginosa* is capable of growth in a wide range of different conditions, including soil and coastal marine habitats or plant and animal tissues (1, 2). *P. aeruginosa* is also a significant opportunistic human pathogen recently described as a species urgently requiring development of novel antibiotics for treatment of disease due to the emergence of multidrug-resistant strains (3). *P. aeruginosa* is able to infect patients suffering from burns, immunosuppression, and cystic fibrosis (CF); reduced pulmonary function caused by chronic *P. aeruginosa* infection is the largest cause of mortality in cystic fibrosis patients (1, 2, 4).

Diversity in cultivation of habitats is likely underpinned by adaptation mechanisms of *P. aeruginosa* to alter phenotypic behavior. This marked pleiotropism identifies a broad array of environmental cues and a number of archetypal bacterial responses. These resulting bacterial responses might include movement away from or toward a specific chemical stimulus, also known as chemotaxis (5). Alteration of gene expression may also be directly induced by a stimulus, often as part of a two-component regulatory system (6). Finally, transition from a planktonic phenotype to a sessile biofilm lifestyle makes *P. aeruginosa* an important human pathogen causing chronic infection (7, 8). This transition, and with it, bacterial virulence, is critically regulated by intracellular c-di-GMP levels (9). *P. aeruginosa*, and particularly the reference strain PAO1, is an extensively studied model organism in biofilm formation (2, 10, 11).

Adaptive responses that require mechanisms of signal perception are often directly transmitted through sensory proteins. A classic, versatile, and very widespread protein architecture used in many sensory proteins is the Per-Arnt-Sim or PAS domain (12, 13). The domain was first identified to be conserved between the fly clock protein PERIOD, the vertebrate aryl hydrocarbon nuclear translocator (ARNT), and the fly developmental regulator single-minded (SIM) (14, 15), and PAS domains are found across all biological kingdoms. They are frequently found among bacterial sensory systems (15) and play crucial roles in environmental responses of *P. aeruginosa* (16, 17). They are also widespread in regulators of intracellular c-di-GMP, where they are suggested to play a role in regulation of virulence (18), as well as motility and biofilm phenotype (19).

Cache domains are the extracellular relatives of the intracellular PAS domains (20–22). Typically acting as signal receptors, they bind small ligands and propagate signals into the cell interior, suggested to be mediated by the C-terminal helix that crosses the membrane (21–24). They are often classified into sCache or dCache domains with one or two PAS-like domains, respectively (22). While Cache domains are the predominant superfamily of extracellular receptors in prokaryotes, they are also found as extracellular domains ubiquitous across all kingdoms (22).

Adaptation of PAS and Cache domains to a variety of signals is achieved by cofactor or ligand binding. Further, they can typically dimerize in response to physiological change and, in doing so, alter the activity of the effector outputs within PAS domain-containing proteins (14, 25). Many proteins contain several PAS domains that may be common in structure but different in function, and hence study of cofactor or ligand binding is a systems approach that is essential to determine how responses are regulated.

Here, we update the list of known PAS and Cache domains in *P. aeruginosa* PAO1, provide a phylogenetic analysis, and add functional insights using a sequence/structure-based approach, with prediction of cofactor or ligand binding. Building upon earlier identification through Hidden Markov Model analysis (22, 26), the analysis includes a total of 90 PAS domains, 2 sCache domains, and 9 dCache domains that were identified from a total of 70 genes. Phylogeny and physiology are put in context, as illustrated here for carboxylic acid-binding PAS domains.

RESULTS

Selection of *P. aeruginosa* PAO1 PAS and Cache domains. The PAS domain fold is a highly conserved yet versatile structure. Sequence identity of PAS domains is typically below 20 percent (14, 15, 27), making the identification of all PAS domains in an organism difficult. Hidden Markov Models (HMM) were used as a sensitive method for

homology detection, and these methods are typically employed for cases with low sequence identity (28). Application of HMM methods through HMM-to-HMM comparison (implemented through HHblits [29]) has initially identified 106 domains from 70 PAO1 proteins (26), of which 18 form 9 dCache domains and 2 were later classified as sCache, thus leaving 86 bona fide PAS domains (22).

The search used here employed domain boundaries predicted from the earlier studies. Additionally, the 70 PAO1 proteins were queried with the SMART domain prediction server (30, 31). The SMART analysis uses different criteria for minimum sequence length compared with the more stringent HMM-to-HMM analysis (26). SMART identifies the shorter PAS and PAC sequence motifs separately (32) and was thus able to identify the N-terminal PAS motif for PA4021, PA4112, PA4147, and PA4197. Sequences were extended at the C terminus to a length of ~120 amino acids to facilitate further analysis. The final list of 101 PAS and Cache domains selected from this analysis is given in Table 1.

Phylogenetic analysis and grouping of PAS and Cache domains. We performed a grouping of sequences by maximum likelihood phylogeny to understand the relationships between sequences (see Materials and Methods). Neighborhood within this phylogenetic analysis might infer similar properties with regard to cofactor or ligand binding. The relationship between proteins grouped within the same clade can thus be used as an indicator toward a functional assignment of the individual domains, leading to experimentally testable hypotheses.

The sequence alignment was performed against a reference data set made up from PAS and Cache domains with known ligand or cofactor obtained from structural analysis. Table 2 indicates cofactor or ligand identified in these structures and size of the ligand- or cofactor-binding pocket. The reference data set also included 37 sequences for ligand- or cofactor-free PAS or Cache structures, selected based on physiological roles not requiring ligand or cofactor, e.g., mediating dimerization or downstream signaling in response to a conformational change in a multidomain protein or as a result of binding to another protein. This ties in with the observation that a significant number of PAS or Cache domains were reported as structures without an associated cofactor or ligand (15).

The maximum likelihood phylogeny analysis with a 100-replicate bootstrap consensus tree is shown in Fig. 1. For the PAS and Cache domains analyzed here, we found that maximum likelihood grouped PAS or Cache domains from the reference data set into clades of similar cofactor or ligand binding across the largest number of bootstrap replicates, in comparison to other ways of constructing phylogenetic trees (see supplemental material for further detail). The phylogenetic tree identifies a number of clades and groups PAO1 sequences together with structurally known PAS and Cache domains. The grouping is based solely on the phylogenetic analysis and is thus unbiased by ligand or cofactor binding or structural knowledge. A number of PAO1 PAS and Cache domains have been characterized previously with respect to ligand or cofactor binding, and the fact that these sequences cluster in the tree with the reference sequence from structures with similar ligand or cofactor validates the approach taken here.

Assignments were made based on the basis of clustering in more than 15 bootstrap replicates, as this threshold provides unambiguous clustering of the reference PAS and Cache domains in almost all cases while retaining the clustering of known homologues. For previously uncharacterized PAO1 PAS or Cache domains, inference suggests that grouping of *P. aeruginosa* sequences with structural representatives from the reference indicates similar ligand or cofactor binding. Alignments of individual clades are presented in the supplemental material, and we give a few examples in the following section.

Inferences from example clades and grouping of PAS and Cache domains. A prominent clade, marked with arrow 1 in Fig. 1 (alignment found in the supplemental material), places the PAO1 PA1336, PA5165, and PA5512 dCache domains with the reference structure sequences of the two DctB dCache domains of *Vibrio cholerae* and

TABLE 1 *P. aeruginosa* PAO1 proteins with PAS or Cache domains[a]

		Domain boundary			
Gene	**Protein**	**PAS1**	**PAS2**	**PAS3**	**PAS4**
PA0172	SiaA	dCache 102–304	dCache 102–304		
PA0176	Aer2/TlpG/McpB	166–287			
PA0285		79–198	206–320		
PA0290		31–151			
PA0338		50–170			
PA0464	CreC	sCache35–179			
PA0533		12–135	137–255	265–379	
PA0575		310–426	438–550	562–675	682–797
PA0600	AgtS	323–436	446–568		
PA0847		142–284	444–560		
PA0861	RbdA	243–363			
PA0873	PhhR	82–187			
PA0928	GacS	43–161			
PA1098	FleS	74–164			
PA1120	TpbB/ YfiN	46–152			
PA1180	PhoQ	33–161			
PA1181	YegE	298–415	427–542	553–674	
PA1196	DdaR	20–132			
PA1243		57–169	343–456		
PA1261	IhpR	1–107			
PA1336	AauS	dCache51–346	dCache51–346		
PA1347		23–129			
PA1423	BdlA	3–112	116–234		
PA1438	MmnS	41–166			
PA1561	Aer/ TlpC	8–121			
PA1611		38–169			
PA1930	McpS	17–134	139–254		
PA1976	ErcS′	97–207	226–338	339–454	
PA1992	ErcS	41–157			
PA2005	HbcR	17–123			
PA2072		301–414			
PA2177		62–180	190–308		
PA2449		79–182			
PA2480		30–148			
PA2524	CzcS	34–171			
PA2652	CtpM	sCache42–198			
PA2654	TlpQ	dCache50–346	dCache50–346		
PA2824	SagS	56–169			
PA2870		97–211	241–348		
PA3044	RocS2	110–225			
PA3271		636–751			
PA3946	RocS1	573–687			
PA4021	EatR	80–185	225–344		
PA4036		432–537			
PA4112		343–460	491–614	626–744	
PA4117	BphP	23–123			
PA4147	AcoR	82–191	225–344		
PA4197	BfiS	158–265	266–383	389–504	
PA4290		411–520			
PA4293	PprA	303–421	431–549	560–675	
PA4307	PctC	dCache34–275	dCache34–275		
PA4309	PctA	dCache35–273	dCache35–273		
PA4310	PctB	dCache35–274	dCache35–274		
PA4398		50–154	286–395		
PA4546	PilS	195–296			
PA4581	RtcR	52–165			
PA4601	MorA	290–411	582–705	717–845	825–967
PA4633		dCache51–346	dCache51–346		
PA4725	CbrA	630–739			
PA4886		69–166			
PA4959	FimX	142–254			

(Continued on next page)

TABLE 1 (Continued)

Gene	Protein	Domain boundary			
		PAS1	PAS2	PAS3	PAS4
PA4961		53–166			
PA4982	AruS	288–388			
PA5017	DipA	9–130	344–460		
PA5124	NtrB	3–116			
PA5165	DctB	dCache44–291	dCache44–291		
PA5361	PhoR	101–201			
PA5442		275–393	401–515		
PA5484	KinB	257–369			
PA5512	MifS	dCache31–298	dCache31–298		

[a]Of the 70 genes listed, several encode more than one PAS domain. Domain boundaries were identified by HMM analysis in previous studies (22, 26) or with the SMART domain web server (30, 31).

Sinorhizobium meliloti in 29 out of 100 bootstrap replicates. While PA5165 has previously been identified as a DctB homologue within *P. aeruginosa* (33, 34), this clade gives new insight, as it implies coevolution with PA1336 and PA5512. It may further predict the potential for binding similar ligands in all five domains.

A close relationship is detected between PA5124 PAS1 and the PAS domain from *Escherichia coli* DhaR, identified in 93 out of 100 bootstrap replicates, marked with arrow 2 in Fig. 1. The *E. coli* DhaR protein is a regulator of transcription. The PAS domain of DhaR contains a very small cavity that precludes binding of larger organic cofactors (Table 2), and the PAS domain is instead thought to be involved in signal transmission through dimerization (35). Conformational changes of the entire protein would be induced by binding of a number of different known interaction partners (35). By inference, the PA5124 PAS1 domain may also not bind any cofactor. This finding may be surprising, as the two proteins have vastly different domain architectures: DhaR consists of a GAF domain, a PAS domain, and a C-terminal domain involved in interaction with σ^{70} (35), while PA5124 is predicted to have a single PAS domain, as well as a histidine kinase and an accompanying phospho-transfer domain known from two-component signaling pathways (6, 30, 31).

The phylogenetic tree in Fig. 1 shows many PAO1 PAS and Cache domains that do not cluster to the chosen reference data set, guided by choice of our reference cofactor- and ligand-binding domain data set. However, the phylogenetic analysis performed here still provides insight into the evolutional origin of several of these domains. An example cluster identified in 96% of replicates is marked with arrow 3 in Fig. 1 and contains the two PAS domains of PA1423 (BdlA) and the two PAS domains of PA1930 (McpS). The protein architectures are similar, as both proteins consist of two PAS domains N-terminal to a methyl-accepting chemotaxis domain. PA1423 and PA1930 are unique within PAO1, as they possess methyl-accepting chemotaxis domains shorter than those of any other chemoreceptors (36). The similar architecture suggests functional differentiation for these proteins.

Assignment of cofactor or ligand binding based on sequence motif. The ability of a PAS or Cache domain to bind cofactor or ligand should be reflected in conservation of cofactor- or ligand-interacting amino acids. PAS and Cache domains are structurally homologous, albeit with overall rather low sequence identity. To add additional information, the alignment can therefore be constrained by predicted secondary structure. Using the combined primary and predicted secondary information then gives sufficient confidence for modeling of the 3D localization of conserved cofactor- or ligand-coordinating amino acids.

We use this approach here to inspect the ligand- or cofactor-binding environment. The PAO1 test data set was aligned to the different cofactor- and ligand-binding subsets of the reference data set, using secondary structure constraints through use of PROMALS3D (37). Detection of conservation or conservative substitution of amino acids known to form sidechain-mediated interactions within the resulting alignments

TABLE 2 The reference data set contains sequences from PAS or Cache domain structures, grouped by physiological cofactor or ligand and by protein and species name, as well as references to the structural database and literature[a]

Cofactor or ligand	M_w of cofactor or ligand (g/mol)	Protein	Organism	PDB	PAS domain boundary from PDB RCSB	Pocket MS vol (Å^3)	Comment
4′-Hydroxycinnamic acid	164.16	Ppr	*Rhodospirillum centenum*	1MZU (70)	25–129	397.0	
	164.16	PYP	*Halorhodospira halophila*	2PHY (58)	1–125	226.5	
Autoinducers	124.14	VqmA	*Vibrio cholerae*	6IDE (39)	16–121	318.0	
Aromatics	92.14	TodS	*Pseudomonas putida*	5HWV (71)	5–133	216.8	
FAD	785.55	MmoS (PAS A)	*Methylococcus capsulatus*	3EWK (72)	1–100	763.2	
	785.55	NifL	*Azotobacter vinelandii*	2GJ3 (57)	16–117	548.1	
	785.55	Vivid	*Neurospora crassa*	2PDR (73)	35–149	691.8	
Fatty acids	228.37	Caur_2278/ MltR	*Chloroflexus aurantiacus*	3PXP (74)	111–292	965.3	
	356.54	HIF3a9 PAS-B	*Homo sapiens*	4WN5 (75)	235–343	1,108.0	
	200.32	RpfR	*Cronobacter turicensis*	6DGG (76)	7–110	453.7	
	256.42	Rv1364c	*Mycobacterium tuberculosis*	3K3C (38)	27–132	615.0	
FMN	456.34	Aureochrome 1a LOV	*Phaeodactylum tricornutum*	5A8B (77)	34–138	552.8	
	456.34	Cagg_3753	*Chloroflexus aggregans*	6RHG (78)	48–152	612.1	
	456.34	LOV	*Dinoroseobacter shibae*	6GAY (79)	32–141	858.9	Pocket open to solvent
	456.34	EI222	*Erythrobacter litoralis*	3P7N (80)	34–141	585.6	
	376.36	EL346 (HTCC2694)	*Erythrobacter litoralis*	4R38 (81)	15–123	625.9	Riboflavin binding
	456.34	Env1	*Hypocrea jecorina*	4WUJ (82)	37–146	589.6	
	456.34	LOV	*Rhodobacter Sphaeroides*	4HIA (83)	18–123	627.8	
	456.34	LOV-HK	*Brucella abortus*	3T50 (56)	26–140	573.7	
	456.34	NPH1-1 (LOV2)	*Avena sativa*	2V0U (84)	13–119	587.7	
	456.34	AUREO1	*Ochromonas danica*	6I20 (85)	16–120	633.0	
	456.34	PAL PAS B	*Nakamurella multipartita*	6HMJ (86)	209–347	669.9	
	456.34	Phot	*Chlamydomonas reinhardtii*	1N9L (87)	17–125	730.3	
	456.34	Phot1	*Arabidopsis thaliana*	2Z6C (88)	15–125	629.9	
	456.34	Phot2	*Arabidopsis thaliana*	2Z6D (88)	16–121	741.6	
	456.34	Phy3	*Adiantum capillus-veneris*	1G28 (89)	929–1032	715.6	
	456.34	SB1-LOV	*Pseudomonas putida*	3SW1 (90)	16–119	1,075.2	Pocket open to solvent
	456.34	AUREO1	*Vaucheria frigida*	3ULF (91)	51–154	581.8	
	456.34	YtvA	*Bacillus subtilis*	2PR5 (25)	8–111	738.5	
	456.34	Ado1 LOV	*Arabidopsis thaliana*	5SVG (92)	16–129	602.0	
Heme-B	616.49	Aer2	*Pseudomonas aeruginosa*	3VOL (93)	32–135	814.2	
	616.49	Aer2	*Vibrio cholerae*	6CEQ (94)	170–280	1,007.5	
	616.49	DosP	*Escherichia coli*	1V9Y (95)	30–132	565.1	
	616.49	HODM	*Pseudomonas mendocina*	5LTE (96)	155–290	1,869.9	
	616.49	FixL	*Bradyrhizobium japonicum*	1DRM (97)	13–117	984.8	
	616.49	FixL	*Rhizobium meliloti*	1D06 (98)	26–130	907.4	

(Continued on next page)

TABLE 2 (Continued)

Cofactor or ligand	M_w of cofactor or ligand (g/mol)	Protein	Organism	PDB	PAS domain boundary from PDB RCSB	Pocket MS vol (Å^3)	Comment
Heme-C	616.49	GSU0582	*Geobacter sulfurreducens*	3B47 (99)	45–131	24.8	Non-classical heme cofactor binding
	616.49	GSU0935	*Geobacter sulfurreducens*	3B42 (99)	45–127	12.1	
	618.50	Tll0287	*Thermosynechococcus elongatus*	5B82 (100)	26–186	1,196.8	Extended pocket
Metals	107.87	CusS	*Escherichia coli*	5KU5 (101)	38–185	58.1	
	65.39	CzcS	*Pseudomonas aeruginosa*	5GPO (102)	38–161	91.5	
No cofactor or ligand binding	NA	Agp1 (Atu1990)	*Agrobacterium fabrum*	5HSQ (103)	20–108	52.7	
	NA	Agp2 (Atu2165)	*Agrobacterium fabrum*	6G1Y (104)	21–119	175.1	Pocket open to solvent
	NA	AhR	*Homo sapiens*	5NJ8 (105)	106–253	67.5	
	NA	AhR	*Mus musculus*	4M4X (106)	41–186	42.3	
	NA	AhRR	*Homo sapiens*	5Y7Y (107)	A102–A256	133.0	
	NA	ARNT (PAS A)	*Bos taurus*	5Y7Y (107)	89–189	1,188.7	Open binding groove
	NA	ARNT (PAS B)	*Bos taurus*	5Y7Y (107)	B208–B311	128.2	
	NA	ARNT (PAS B)	*Homo sapiens*	1X0O (108)	1–119	38.9	
	NA	ARNT (PAS A)	*Mus musculus*	4ZP4 (109)	92–263	48.1	
	NA	ARNT (PAS B)	*Mus musculus*	4ZP4 (109)	282–384	137.3	
	NA	BMAL1/ARNTL (PASB)	*Mus musculus*	4F3L (110)	277–382	234.3	Pocket w/o occupancy
	NA	CLOCK (PAS B)	*Mus musculus*	4F3L (110)	250–353	146.5	
	NA	Cph1	*Synechocystis sp.*	2VEA (111)	29–126	65.8	
	NA	DhaR/ YcgU	*Escherichia coli*	4LRX (35)	C214–C305	90.7	
	NA	BphP	*Deinococcus radiodurans*	1ZTU (112)	52–144	79.7	
	NA	EAG/ Kcnh1	*Mus musculus*	4LLO (113)	B23–B134	112.5	
	NA	EAG/ Kcnh1	*Rattus norvegicus*	5K7L (114)	A27–A132	186.6	
	NA	PadC	*Idiomarina species A28L*	5LLW (115)	33–123	86.7	
	NA	MmoS (PAS B)	*Methylococcus capsulatus*	3EWK (72)	122–227	88.0	
	NA	NcoA1 PAS B	*Homo sapiens*	5NWM (116)	A254–A385	410.7	
	NA	NcoA-1/ SRC-1	*Mus musculus*	1OJ5 (117)	A259–A367	204.5	Pocket w/o occupancy
	NA	BphP	*Pseudomonas aeruginosa*	3C2W (118)	25–114	136.1	
	NA	PhyB	*Arabidopsis thaliana*	4OUR (119)	29–131	41.1	
	NA	PpsR (N-PAS)	*Rhodobacter sphaeroides*	4HH2 (120)	29–125	25.7	
	NA	PpsR (PAS1)	*Rhodobacter sphaeroides*	4HH2 (120)	166–261	23.4	
	NA	PpsR (PAS2)	*Rhodobacter sphaeroides*	4HH2 (120)	284–383	56.3	
	NA	BphP1 PAS1	*Rhodopseudomonas palustris*	4GW9 (121)	54–145	53.2	
	NA	BphP1 PAS2	*Rhodopseudomonas palustris*	4GW9 (121)	549–646	121.2	
	NA	BphP2	*Rhodopseudomonas palustris*	4E04 (122)	29–121	21.0	
	NA	BphP3	*Rhodopseudomonas palustris*	2OOL (123)	42–138	100.1	
	NA	BphP	*Stigmatella aurantiaca*	6BAF (124)	17–112	12.3	
	NA	BphP2	*Stigmatella aurantiaca*	6PTQ (125)	19–103	31.4	

(Continued on next page)

TABLE 2 (Continued)

Cofactor or ligand	M_w of cofactor or ligand (g/mol)	Protein	Organism	PDB	PAS domain boundary from PDB RCSB	Pocket MS vol (Å^3)	Comment
	NA	Soluble guanylate cyclase (sGC) PAS α domain	*Manduca sexta*	4GJ4 (126)	10–110	31.4	
	NA	Soluble guanylate cyclase (sGC) α subunit	*Homo sapiens*	6JT0 (127)	A288–A386	47.1	
	NA	Soluble guanylate cyclase (sGC) β subunit	*Homo sapiens*	6JT0 (127)	B217–B326	72.5	
	NA	XccBphP (N-terminal PAS domain)	*Xanthomonas campestris*	5AKP (128)	33–128	97.4	
	NA	XccBphP (C-terminal PAS domain)	*Xanthomonas campestris*	5AKP (128)	534–637	852.6	Open binding groove
dCache - amino acids	89.09	CtaA	*Pseudomonas fluorescens*	6PXY (129)	41–269	147.3	
	89.09	Mlp24/ McpX/ VC_A0923	*Vibrio cholerae*	3C8C (21)	1–226	138.5	
	105.09	Mlp37	*Vibrio cholerae*	5AVE (130)	5–234	127.5	
	149.21	PctA	*P. aeruginosa*	5LTX (44)	29–256	334.5	
	175.21	PctB	*P. aeruginosa*	5LT9 (44)	33–256	227.8	
	103.12	PctC	*P. aeruginosa*	5LTV (44)	33–257	191.8	
	115.13	PscC	*Pseudomonas syringae*	6MNI	23–275	192.5	
	131.17	Tlp3	*Campylobacter jejuni*	4XMR (131)	37–285	248.9	
	111.14	TlpQ	*P. aeruginosa*	6FU4 (132)	39–323	753.7	Open binding groove
dCache - cytosine	111.10	Dret_0059	*Desulfohalobium retbaense*	5ERE (133)	322–562	349.1	
dCache - phosphate	94.97	VP0354 (vpHK1S-Z8)	*Vibrio parahaemolyticus*	3LID (21)	8–269	224.1	
dCache - polyamines	88.15	McpU	*Pseudomonas putida*	6F9G (134)	41–300	615.0	
Cache - no cofactor or ligand binding	NA	LuxQ	*Vibrio cholerae*	3C38 (21)	21–240	37.0	
	NA	LuxQ	*Vibrio harveyi*	2HJE (135)	2–221	23.7	
dCache - QACs	144.19	McpX	*Rhizobium meliloti*	6D8V (136)	38–306	229.7	
dCache - cytokinins	203.24	AHK4	*Arabidopsis thaliana*	3T4J (137)	126–393	528.2	
dCache - carboxylic acids	118.09	DctB	*Rhizobium meliloti*	3E4O (138)	48–301	134.1	
	118.09	DctB	*Vibrio cholerae*	3BY9 (24)	27–285	130.3	
	88.06	KinD	*Bacillus subtilis*	4JGO (139)	6–204	156.2	
	90.08	TlpC	*Helicobacter pylori*	5WBF (140)	3–261	263.2	
sCache - acetate sensing	59.04	Adeh_3718	*Anaeromyxobacter dehalogenans*	4K08 (47)	57–144	84.9	
sCache - carboxylic acids	189.10	CitA	*Klebsiella pneumoniae*	1P0Z (50)	50–126	356.9	Pocket open to solvent
	134.09	DcuS	*Escherichia coli*	3BY8 (24)	56–130	174.9	

(Continued on next page)

TABLE 2 (Continued)

Cofactor or ligand	M_w of cofactor or ligand (g/mol)	Protein	Organism	PDB	PAS domain boundary from PDB RCSB	Pocket MS vol (Å^3)	Comment
	88.06	VP0183	*Vibrio parahaemolyticus*	4EXO (45)	56–146	95.9	
	73.07	PscD-SD	*Pseudomonas syringae*	5G4Y (46)	32–178	98.7	
sCache - metals	58.69	PhoQ	*Escherichia coli*	3BQ8 (59)	41–138	108.5	
	40.08	PhoQ	*Salmonella enterica* serovar Typhimurium	1YAX (23)	39–138	32.8	
sCache - urea	60.05	TlpB	*Helicobacter pylori*	3UB6 (45)	70–156	170.9	

[a]PAS or Cache domain boundaries are indicated. The pocket or cavity volume is presented along with the molecular weight (M_w) of the cofactor or ligand in the pocket/cavity, where present. MS, pocket volume based on the molecular surface; QAC, quaternary ammonium compound.

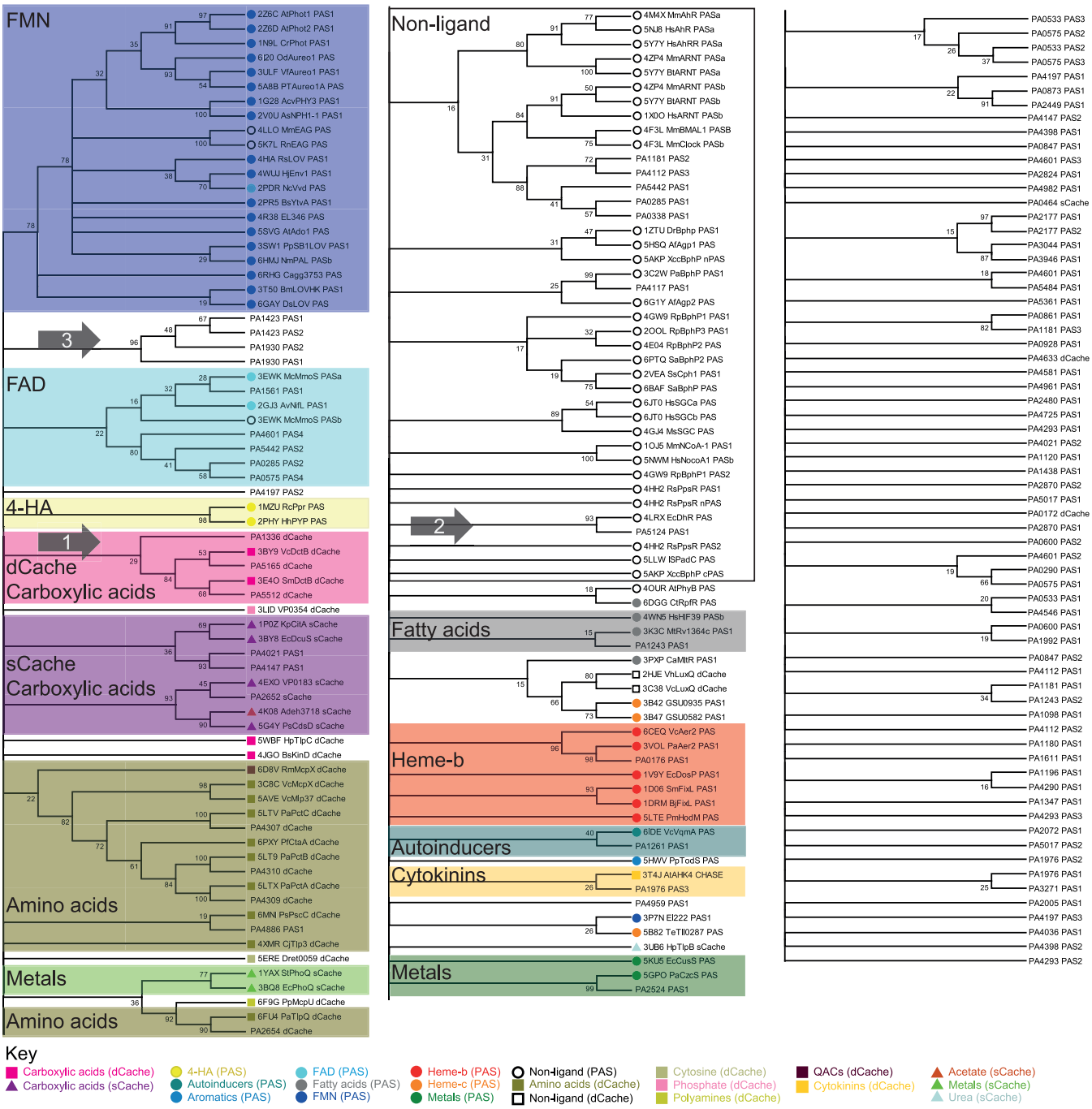

FIG 1 Maximum likelihood phylogenetic analysis of *Pseudomonas aeruginosa* PAO1 PAS or Cache domains with the reference set of structurally characterized domains. The percentage of bootstrap replicates that reproduced each branch is given, with branches corresponding to less than 15% of bootstrap replicates collapsed and rearranged for clarity. PAS, dCache, and sCache domains are labeled with a circle, square, or triangle, respectively. The nature of ligand or cofactor is given in the key and denoted by color, and individual alignments of these groups are found in the supplemental material. Groups discussed in the text are marked with a numbered arrow. The supplement to this article contains an evaluation of different phylogenetic analyses and alignments of individual clades shown in Fig. 1 and discussed in the text.

was taken to be indicative of the capacity to bind cofactor or ligand. Cofactor or ligand binding capacity was thus inferred from phylogenetic analysis and conservation of key amino acid residues.

The dCache domains of PA1336, PA5165, and PA5512 that were identified to group with sequences of carboxylic acid-binding dCache domain structures from the reference data set (arrow 1 in Fig. 1) were investigated to confirm conservation of ligand-coordinating residues. Amino acids responsible for substrate coordination within the carboxylic acid-binding

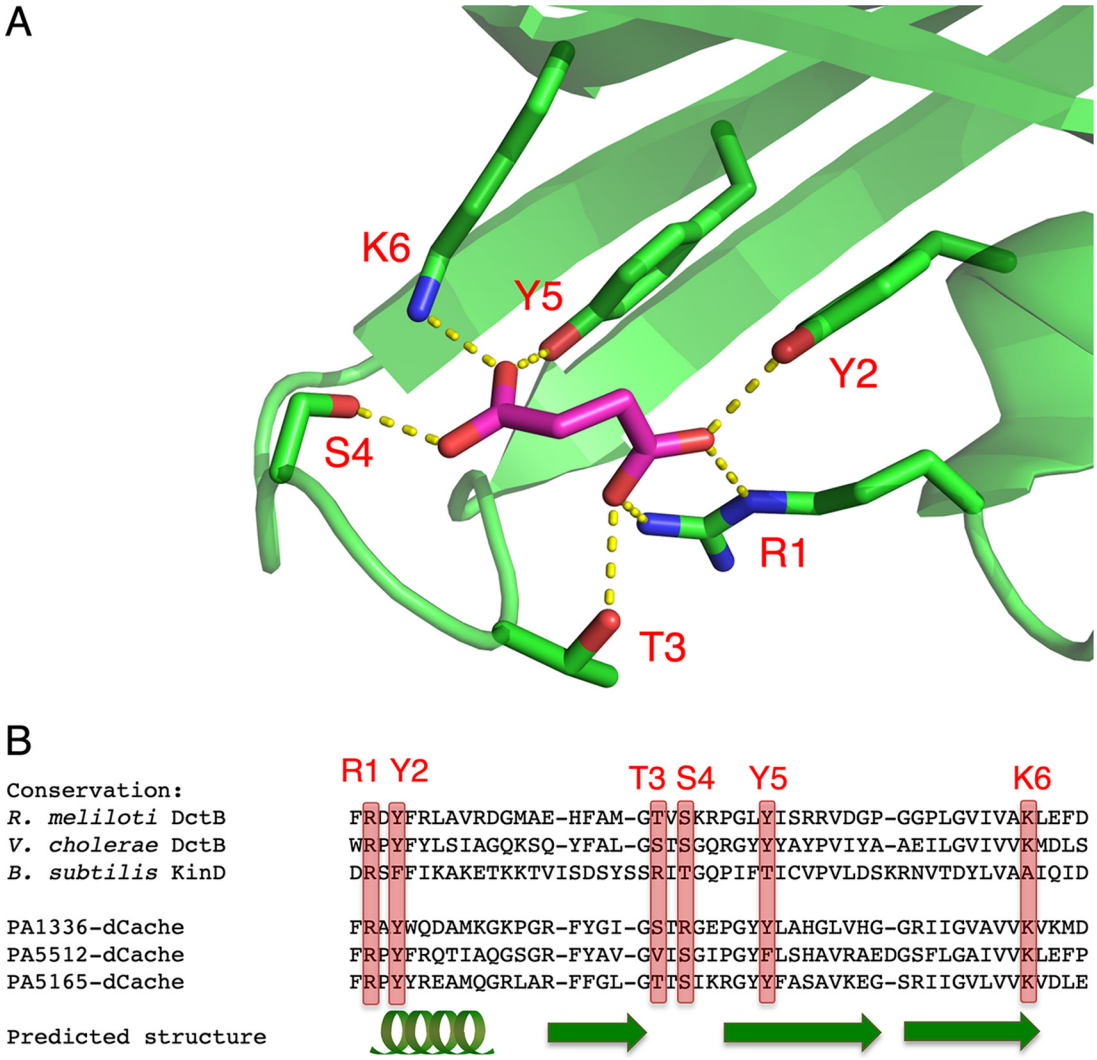

FIG 2 Ligand-binding capacity of the dCache phylogenetic clade investigated by primary/secondary structure conservation analysis. Sequences selected here are highlighted by arrow 1 in Fig. 1. (A) The dCache domain of *Rhizobium meliloti* DctB in complex with succinate (PDB 3E4O). Residues involved in the coordination of succinate are shown as sticks and are labeled with single-letter amino acid codes and consecutive numbers. (B) Guided sequence alignment using the predicted secondary structure for PAO1 PA1336, PA5165, and PA5512 against the carboxylic acid-binding dCache domains from *R. meliloti* DctB, *Vibrio cholerae* DctB, and *Bacillus subtilis* KinD. The predicted secondary structure used for alignment is denoted as a cartoon under the sequences. The position of amino acids used for ligand binding in DctB is indicated in the alignment.

dCache domain of DctB are shown in Fig. 2. It is worth pointing out that these residues vary between KinD carboxylic acid-binding dCache domains, where substrates (succinate, malonate, pyruvate, and lactate) adopt a different binding pose, and DcuS sCache domains of different organisms that display different substrate specificity (citrate, malate, pyruvate, and propionate). PA5512, PA1336, and PA5165 use the same repertoire of ligand-coordinating amino acid residues as DctB. However, the slight variation needs to be discussed with respect to substrate specificity.

An interesting example evaluates fatty acid-binding PAS domains, of which a number of different binding poses and interactions have been structurally characterized (alignment found in supplemental material). The reference protein Rv1364c from *Mycobacterium tuberculosis* contains a binding motif used to coordinate palmitic acid in reference 38. PA0847 PAS2, PA1196 PAS1, PA1976 PAS3, and PA4112 PAS2 all display conservation of the two relevant amino acids responsible for side chain-specific ligand interaction. In PA1196, there is a conservative exchange (aspartate to glutamate)

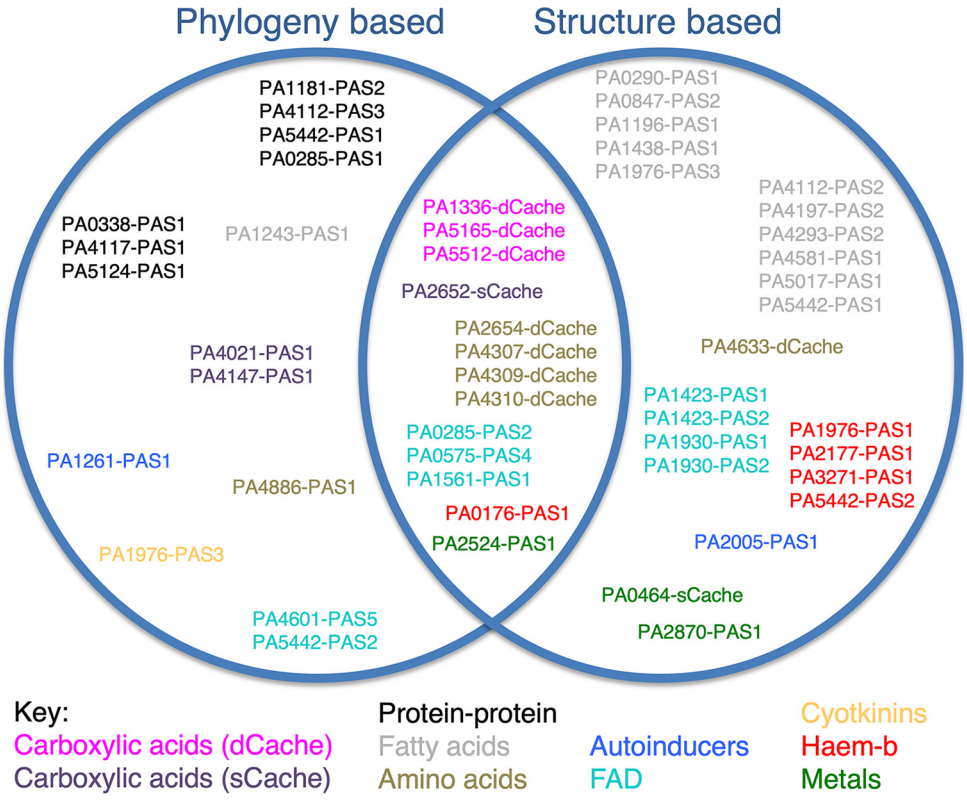

FIG 3 Combination of phylogeny- and conservation-based assignment. *Pseudomonas aeruginosa* PAS and Cache domains predicted to bind cofactors or ligands are grouped by method of prediction. The nature of the bound cofactor or ligand is denoted by color.

in one of these two recognition amino acids. The analysis of conservation performed here adds significantly to the phylogenetic analysis, as the majority of fatty acid-binding PAS domains were assigned on the basis of conserved ligand-binding amino acids.

Indeed, a number of novel assignments can be made based on conserved binding motif. For example, PA2005 PAS1 could not be placed into a clade through phylogenetic analysis but is assigned here as autoinducer binding based on conservation of the two residues critical for side chain-specific coordination of the autoinducer DPO (3,5-dimethylpyrazin-2-ol), seen in *V. cholerae* VqmA (6IDE) (39). The alignment reveals that one amino acid is conserved while the other one is a conservative exchange from lysine to arginine. The data generated are summarized in Fig. 3 and Table 3.

DISCUSSION

Individual domains are the building blocks of modular proteins and are required for functional diversification of the proteome. Understanding of protein function is crucially dependent on our grasp of physiological and functional roles of these constituting domains. For the omnipresent PAS and Cache domains, analysis is generally hampered by failure to predict cofactor- or ligand-binding state from sequence. Even identification of these domains proves to be difficult, due to low sequence conservation. While HMM sequence searches are a sensitive method to detect homology in cases of low sequence identity, we also made use of the SMART domain prediction server (30, 31) and used structure-guided analyses here.

We have studied the bacterial model organism *Pseudomonas aeruginosa* that has the ability to adapt to various environmental conditions, a survival strategy and an underling property important in the clinical setting. We sought to identify the nature of cofactors and ligands that bind to PAS or Cache domains within *P. aeruginosa* PAO1 using phylogeny and structural conservation analyses. A number of differences exist

TABLE 3 PAS or Cache domains and predicted cofactors or ligands assigned on the basis of combined phylogeny and sequence-structure alignment

Cofactor or ligand	Protein	Domain	Known physiological role
Amino acids	PA2654 (TlpQ)	dCache	Chemotaxis toward ethylene and histamine (132, 141, 142)
	PA4307 (PctC)	dCache	Chemotaxis toward amino acids (44, 143, 144)
	PA4309 (PctA)	dCache	Chemotaxis toward amino acids (44, 143–145)
	PA4310 (PctB)	dCache	Chemotaxis toward amino acids (44, 143, 144)
	PA4633	dCache	Unknown (146)
	PA4886	PAS1	Unknown (147, 148)
Autoinducers	PA1261 (LhpR)	PAS1	Transcriptional regulator (149)
	PA2005 (HbcR)	PAS1	Regulation of (R)-3-hydroxybutyrate catabolism (150)
Carboxylic acids -dCache like	PA1336 (AauS)	dCache	Regulation of genes involved in aspartate, glutamate, and glutamine uptake and catabolism (43)
	PA5165 (DctB)	dCache	Regulation of C_4-dicarboxylic acid transport systems (34)
	PA5512 (MifS)	dCache	Regulation of α-ketoglutarate transport and utilization (41, 42)
Carboxylic acids -sCache like	PA2652 (CtpM)	sCache	Chemotaxis toward malate (48, 49, 151)
	PA4021 (EatR)	PAS1	Regulation of ethanolamine catabolism (51)
	PA4147 (AcoR)	PAS1	Regulation of 2,3-butanediol and acetoin metabolism (52, 53)
Cytokinins	PA1976 (ErcS')	PAS3	Regulation of ethanol oxidation (152, 153)
FAD	PA0285	PAS2	Regulation of biofilm formation (154)
	PA0575	PAS4	Regulation of biofilm formation in response to L-arginine
	PA1423 (BdlA)	PAS1	Regulation of biofilm dispersal (17, 155, 156)
	PA1423 (BdlA)	PAS2	Regulation of biofilm dispersal (17, 155, 156)
	PA1561 (Aer/TlpC)	PAS1	Aerotaxis (157, 158)
	PA1930 (McpS)	PAS1	Regulation of chemotaxis (40)
	PA1930 (McpS)	PAS2	Regulation of chemotaxis (40)
	PA4601 (MorA)	PAS4	Regulation of flagellar development and protease secretion (159–162)
	PA5442	PAS2	Unknown
Fatty acids	PA0290	PAS1	Regulation of biofilm formation and Psl production (154, 163–165)
	PA0847	PAS2	Regulation of motility in response to a no. of stimuli (165, 166)
	PA1196 (DdaR)	PAS1	Regulation of methylarginine metabolism, role in quorum-sensing (167, 168)
	PA1243	PAS1	Regulation of swimming and biofilm formation (169)
	PA1438 (MmnS)	PAS1	Regulation of efflux pump expression (170)
	PA1976 (ErcS')	PAS2	Regulates ethanol oxidation (152, 153)
	PA4112	PAS2	Histidine kinase of unknown pathway
	PA4197 (BfiS)	PAS2	Regulation of biofilm formation (171–174)
	PA4293 (PprA)	PAS2	Regulation of outer membrane permeability/of biofilm formation (175–177)
	PA4581 (RtcR)	PAS1	Homologous to *E. coli* regulator of RNA 3′-terminal phosphate cyclase expression (178–180)
	PA5017 (DipA)	PAS1	Biofilm regulation, chemotaxis, motility, maintenance of c-di-GMP heterogeneity (19, 181–183)
	PA5442	PAS1	Unknown
Heme-b	PA0176 (Aer2/TlpG/McpB)	PAS1	Aerotaxis and virulence (93, 184, 185)
	PA1976 (ErcS')	PAS1	Regulates ethanol oxidation (152, 153)
	PA2177	PAS1	Unknown
	PA3271 (MxtR)	PAS1	Redox sensing and interbacterial signaling (186, 187)
	PA5442	PAS2	Unknown
Metals	PA0464 (CreC)	sCache	Regulation of carbon source catabolism (188, 189)
	PA2524 (CzcS)	PAS1	Regulation of metal detoxification and resistance to carbapenem antibiotics (102, 190–192)
	PA2870	PAS1	Diguanylate cyclase involved in biofilm production, Psl production, regulation of swimming motility (165)
No cofactor or ligand binding	PA0285	PAS1	Regulation of biofilm formation (154)
	PA0338	PAS1	Regulation of biofilm formation, Psl production, and swimming motility (165)
	PA1181 (YegE)	PAS2	Biofilm dispersal (18, 193)
	PA4112	PAS3	Histidine kinase of unknown pathway
	PA4117 (BphP)	PAS1	Quorum sensing (118, 194, 195)
	PA5124 (NtrB)	PAS1	Regulation of nitrogen metabolism, rhamnolipid production, biofilm formation, expression of virulence genes, and swarming (196–200)
	PA5442	PAS1	Unknown

between predictions based on maximum likelihood phylogeny and the individual alignment and inspection of conservation of cofactor- or ligand-interacting amino acids. The PAS domains of PA0873 and PA2449 provide an example in which phylogenetic analysis places them with the reference 4-hydroxycinnamic acid-binding PAS domains. However, when conserved ligand- or cofactor-interacting sidechains were assessed, this classification did not hold. Therefore, analysis based on one method alone may be indicative but not conclusive. The results of our combined analysis and predictions are summarized in Tables 1 and 3, and Fig. 3 highlights differences in assignment from the two different approaches used here.

Our analysis revealed a number of relationships and provides new insight. An example are the four PAS domains marked with the black arrow 3 in Fig. 1 that mark the PAS domains of PA1423 (BdlA) and PA1930 (McpS). Both proteins possess the same architecture, with two N-terminal PAS domains coupled to a methyl-accepting chemotaxis domain. Though they are clearly related, the question of functional diversification arises. Indeed, PA1930 has been reported to have a negative effect on chemotaxis (40), while PA1423 is involved in biofilm dispersal (17). It is therefore likely that the two proteins respond to different triggers and, in doing so, lead to a different biological output. We have experimentally characterized a similar example previously with the proteins PA2072 and RbdA that share an architecture but are responsible for two almost orthogonal functions (19). Thus, there are examples where gene duplication allows proteins to diversify to functionally evolve.

An interesting observation is made here with flavin binders. The distinct clade with sequences of the flavin mononucleotide (FMN)-binding PAS domain structures does not reveal PAO1 PAS domain relatives. However, FAD-binding PAS domains are identified in PAO1. Whether, indeed, FMN is not used as a cofactor in PAO1 remains to be seen. It might turn out that PAO1 has some remarkable and truly distinct PAS domains, and further structural analyses rather than predictions will in time reveal this.

The strength of the combined approach to analyze both phylogeny and conservation of cofactor- or ligand-specifying amino acids is exemplified here with the analysis of carboxylic acid-binding PAS or Cache domains. When the *in silico* results are placed into physiological context, additional insight is gained. We identified carboxylic acid-binding domains in two distinct classes for both sCache and dCache sensory architectures. Variation in substrates and their coordination is detectable between the two classes (24), and consequently, we identify different clades likely to present different substrate interaction and selectivity. The dCache domains illustrate the approach taken by combining sequence, phylogenetic, and structural information. As such, the dCache domains of PA1336, PA5165, and PA5512 are all inferred here to be able to bind carboxylic acids (Fig. 2).

The dCache domain of PA5165 (DctB) was assigned as carboxylic acid binding and, within *P. aeruginosa*, DctB acts as sensor of a two-component pathway involved in regulating the expression of C_4-dicarboxylic acid transport systems (34). It therefore follows that binding of carboxylic acids to the dCache domain of DctB could directly couple levels of C_4-dicarboxylic acids to a signaling cascade responsible for the expression of transport systems used in their uptake (Fig. 4)

Similarly, a carboxylic acid-binding dCache domain within PA5512 (MifS) could also directly link signal perception to a known phenotype (Fig. 4). MifS is required for the transport and utilization of the C_5-dicarboxylic acid α-ketoglutarate (41, 42). It can therefore be hypothesized that this might reveal a potential substrate not previously recognized for dCache domains. As carboxylic acid-binding dCache domains are known to bind C_3 and C_4 substrates with at least one carboxylic acid (15), it is conceivable that the dCache domain in MifS may also be able to bind α-ketoglutarate and act as a sensor. The variation of amino acids identified in the binding pocket from sequence alignments may reflect the required level of flexibility to accept various substrates or binding poses across the variety of PAS and Cache domains in these proteins.

The assignment of PA1336-dCache as carboxylic acid binding could help to identify a source of selectivity within dCache domains. Studies of a protein orthologous to

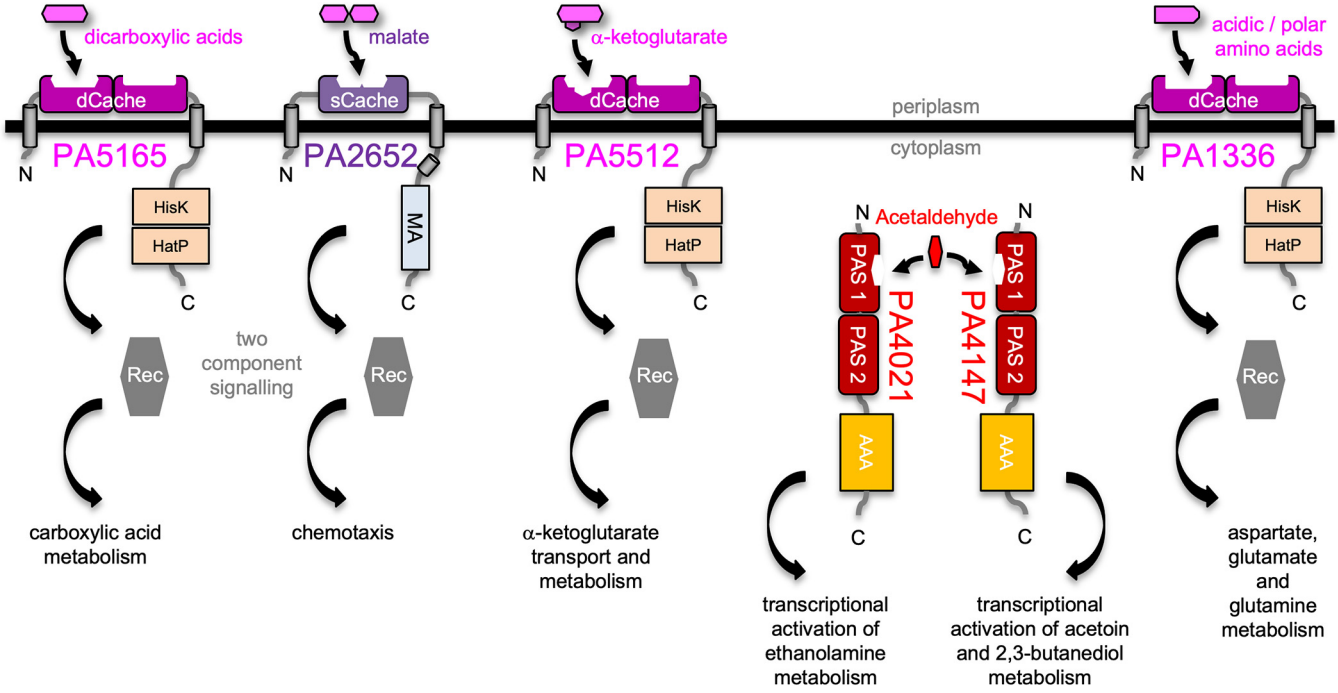

FIG 4 Proteins assigned as containing carboxylic acid-binding Cache and PAS domains are involved in various signaling cascades. PA5165 and PA2652 bind carboxylic acids with periplasmatic dCache/sCache domains. PA5512 is involved in the transport and metabolism of α-ketoglutarate, a previously undescribed ligand for carboxylic acid-binding dCache domains. From the analysis presented here, PA1336 is predicted to bind polar and acidic amino acids. PA5165, PA5512, and PA1336 are sensor histidine kinases (HisK, His kinase A; HatP, histidine kinase-like ATPase), while PA2652 is a chemoreceptor (MA, methyl-accepting). Another cascade in response to acetaldehyde promotes transcriptional changes through interaction with cytoplasmic PAS domains in PA4021 or PA4147; this ligand has not previously been described for PAS domains.

PA1336 (AauS) within *Pseudomonas putida* demonstrate a role in utilization of the amino acids aspartate, glutamate, and glutamine (43), all of which contain side chain carboxylic acid or C=O groups. As PA1336-dCache was classified here as carboxylic acid binding and not amino acid binding, it could be speculated that a similar role for PA1336 within PAO1 to its orthologue in *P. putida* could be accommodated through interaction between the side chains of these amino acids and the PA1136-dCache domain, which then induces a conformational change that activates the two-component system partner of PA1336 to alter gene expression (Fig. 4) (43). Indeed, the highly conserved region identified for amino acid-binding dCache domains (44) is different in PA1336-dCache and instead shows similarity to carboxylic acid-binding dCache domains, with carboxyl groups likely neutralizing the charge through conservation of positively charged side chains of the amino acids labeled R1 and K6 in Fig. 2. These observations may guide future predictions of amino acid selectivity.

Another example of a straightforward link between our analysis and a previously defined physiological function would be the coupling of malate binding to PA2652 (CtpM) with chemotaxis (Fig. 4). The sCache domain of CtpM is assigned as carboxylic acid binding on the basis of its phylogenetic relationship with *Vibrio paraheamolyticus* VP0183 (4EXO) (45), *P. syringae* PscD (5G4Y) (46), and *Anaeromyxobacter dehalogenans* Adeh_3718 (4K08) (47) (Fig. 1) and conservation of five ligand-coordinating amino acids. A function in carboxylic acid binding aligns well with previous reports that CtpM is involved in chemotaxis and has substrate specificity toward malate, which is a known substrate for carboxylic acid-binding sCache domains (24, 48, 49). It may therefore follow from our analysis that the binding of malate to an sCache within CtpM directly couples malate concentration to associated chemotaxis signaling.

Interestingly, phylogeny analysis groups the first PAS domain of PA4021 (EatR) and the first PAS domain of PA4147 (AcoR) with the carboxylic acid-binding sCache domains of *E. coli* DcuS (3BY8) (24) and *Klebsiella pneumoniae* CitA (1P0Z) (50) and the

phosphate-binding dCache domain of *Vibrio parahaemolyticus* VP0354 (3LID) (21). Both PA4021 and PA4147 have known functions as transcriptional regulators for the metabolism of small, hydroxyl-containing, organic compounds and are proposed to perform those functions in response to acetaldehyde (51–53). As acetaldehyde is, to some extent, similar in structure to the carboxylic acids detected by the reference structures in these clades, it is possible that acetaldehyde binding directly to the PAS domains present in PA4021 and PA4147 could form a concise way to induce these changes in transcription and could be the basis of a novel class of PAS domain ligand (Fig. 4).

In conclusion, this study uses protein sequence comparison, phylogeny, and structure-based prediction of ligand or cofactor binding for PAO1 PAS and Cache domains. Although just predictions, the classifications presented give insight from comparison with similar proteins, leading to experimentally testable hypotheses to gain functional insights.

MATERIALS AND METHODS

Selection of *P. aeruginosa* PAO1 PAS and Cache domains. HMM-to-HMM comparisons have previously identified 70 proteins within *P. aeruginosa* PAO1 that contain PAS and Cache domains (22, 26). Protein sequences of these proteins were retrieved from the *Pseudomonas* genome database (11). Selection of the final data set of 101 sequences, containing 91 PAS domains, 9 dCache domains, and 2 sCache domains, is described in Results. These are listed in Table 1.

Generation of the reference data set with 3D structures of PAS and Cache domains. The DALI webserver (54) was used for an exhaustive search of PAS and Cache domains within the Protein Data Bank (PDB; March 2020). Search models were chosen to represent different cofactor- or ligand-binding architectures. PAS domains from *Bradyrhizobium japonicum* FixL (heme-b binder, PDB: 1xj2 [55]), *Brucella abortus* LOV-HK (FMN binder, 3t50 [56]), *Azotobacter vinelandii* NifL (FAD binder, 2gj3 [57]), and *H. halophila* PYP (4′-hydroxycinnamic acid binder, 2phy [58]) were used. Further, sCache domains PhoQ (cation binder, 3bq8 [59]) and DcuS (carboxylic acid binder, 3by8 [24]) from *E. coli* were used. Finally, the dCache domains DctB (carboxylic acid binder, 3by9 [24]) from *V. cholerae* and PctB (amino acid binder, 5lt9 [44]) from *P. aeruginosa* were used. The structures were submitted individually and together retrieved a total of 7,513 matches, corresponding to 986 individual PDB entries. The results included structures not classified as either PAS or Cache domains that were discarded, for example, structurally related GAF domains. Retained were structures with a functional cofactor or ligand bound as well as structures with a reported signaling function, referred to hereafter as "no cofactor or ligand" binding. The final reference data set contained a total of 106 PAS and Cache domains trimmed down to the PAS and Cache domain boundaries and included 78 PAS domains, 20 dCache domains, and 8 sCache domains.

Maximum likelihood phylogeny. The 106 sequences of the reference data set and the 101 sequences from *P. aeruginosa* PAO1 were aligned using CLUSTALW, as implemented in MEGA7 (60, 61). This alignment was then subjected to molecular phylogenetic analysis by maximum likelihood methods within MEGA7 (61, 62). Initial phylogenetic trees were obtained by applying Neighbor-Join (63) and BioNJ (64) algorithms to a matrix of pairwise distances estimated using the JJT-matrix based model (65). The trees were scored and automatically selected based on log-likelihood scores. The bootstrap consensus tree is inferred from 100 replicates and taken to represent the evolutionary history of taxa analyzed (66).

Sequence-structure analysis. The *P. aeruginosa* sCache and dCache domains were aligned to the equivalent subsets in the reference data set using PROMALS3D (37) to determine conservation of ligand-coordinating amino acids residues. Conservation was used to suggest their potential ligand-binding class. For the larger set of PAS domain sequences, the reference data set was divided up according to ligand or cofactor (see Table 2) and then aligned against the *P. aeruginosa* PAS sequences.

Determination of binding pocket or cavity size. The sizes of enclosed cavities or of binding pockets that are open to the surrounding environment allow different classes of PAS or Cache domains to be distinguished. To map their size, the coordinates of the reference structures were uploaded to the CASTp server (67), with the PAS or Cache domain boundaries as given in Table 2. CASTp returns multiple pockets and cavities, which were inspected using UCSF Chimera (68); where a cofactor or ligand was present, this pocket was chosen, but when no cofactor or ligand was identified, the one closest to the center of the PAS or Cache domain was reported, ensuring all cavities/pockets reported here were in a similar position. Table 2 reports the solvent-excluded volume calculated with a probe sphere radius of 1.4 Å, based on Connolly's molecular surface calculation (69), as this parameter was able to discriminate ligand- or cofactor-binding pockets/cavities most clearly. Where the open pockets reported volumes that include not only the actual ligand or cofactor cavity but also the access to the cavity, this is noted in Table 2.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 7.4 MB.

ACKNOWLEDGMENTS

We acknowledge funding by Diamond Light Source and the University of Southampton to A.H. and C.C. Conceptualization, I.T., M.A.W.; investigation, A.H., C.C.; formal analysis, data curation, and validation, A.H., C.C., I.T.; visualization, A.H., I.T.; supervision, M.A.W., J.S.W., I.T.; writing – review and editing, A.H., I.T.; writing – original draft, A.H.; funding acquisition, M.A.W., J.S.W., I.T.

REFERENCES

1. Hardalo C, Edberg SC. 1997. *Pseudomonas aeruginosa* assessment of risk from drinking water. Crit Rev Microbiol 23:47–75. https://doi.org/10.3109/10408419709115130.
2. Stover CK, Pham XQ, Erwin aL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrock-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, Olson MV. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature 406:959–964. https://doi.org/10.1038/35023079.
3. World Health Organization. 2017. WHO publishes list of bacteria for which new antibiotics are urgently needed. World Health Organization, Geneva, Switzerland. http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/.
4. Lyczak JB, Cannon CL, Pier GB. 2002. Lung infections associated with cystic fibrosis. Clin Microbiol Rev 15:194–222. https://doi.org/10.1128/CMR.15.2.194-222.2002.
5. Sampedro I, Parales RE, Krell T, Hill JE. 2015. *Pseudomonas* chemotaxis. FEMS Microbiol Rev 39:17–46. https://doi.org/10.1111/1574-6976.12081.
6. Zschiedrich CP, Keidel V, Szurmant H. 2016. Molecular mechanisms of two-component signal transduction. J Mol Biol 428:3752–3775. https://doi.org/10.1016/j.jmb.2016.08.003.
7. Donlan RM, Costerton JW. 2002. Biofilms: survival mechanisms of clinically relevant microorganisms. Clin Microbiol Rev 15:167–193. https://doi.org/10.1128/CMR.15.2.167-193.2002.
8. Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. 2016. Biofilms: an emergent form of bacterial life. Nat Rev Microbiol 14:563–575. https://doi.org/10.1038/nrmicro.2016.94.
9. Romling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev 77:1–52. https://doi.org/10.1128/MMBR.00043-12.
10. Klockgether J, Munder A, Neugebauer J, Davenport CF, Stanke F, Larbig KD, Heeb S, Schöck U, Pohl TM, Wiehlmann L, Tümmler B. 2010. Genome diversity of *Pseudomonas aeruginosa* PAO1 laboratory strains. J Bacteriol 192:1113–1121. https://doi.org/10.1128/JB.01515-09.
11. Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman FSL. 2016. Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the Pseudomonas genome database. Nucleic Acids Res 44:D646–D653. https://doi.org/10.1093/nar/gkv1227.
12. Zhulin IB, Taylor BL, Dixon R. 1997. PAS domain S-boxes in *Archaea*, *Bacteria* and sensors for oxygen and redox. Trends Biochem Sci 22:331–333. https://doi.org/10.1016/s0968-0004(97)01110-9.
13. Ponting CP, Aravind L. 1997. PAS: a multifunctional domain family comes to light. Curr Biol 7:R674-7. https://doi.org/10.1016/s0960-9822(06)00352-6.
14. Möglich A, Ayers RA, Moffat K. 2009. Structure and signaling mechanism of Per-ARNT-Sim domains. Structure 17:1282–1294. https://doi.org/10.1016/j.str.2009.08.011.
15. Henry JT, Crosson S. 2011. Ligand-binding PAS domains in a genomic, cellular, and structural context. Annu Rev Microbiol 65:261–286. https://doi.org/10.1146/annurev-micro-121809-151631.
16. Huang B, Whitchurch CB, Mattick JS. 2003. FimX, a multidomain protein connecting environmental signals to twitching motility in *Pseudomonas aeruginosa*. J Bacteriol 185:7068–7076. https://doi.org/10.1128/JB.185.24.7068-7076.2003.
17. Morgan R, Kohn S, Hwang S-H, Hassett DJ, Sauer K. 2006. BdlA, a chemotaxis regulator essential for biofilm dispersion in *Pseudomonas aeruginosa*. J Bacteriol 188:7335–7343. https://doi.org/10.1128/JB.00599-06.
18. Kulasakara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN, Hyodo M, Hayakawa Y, Ausubel FM, Lory S. 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. Proc Natl Acad Sci U S A 103:2839–2844. https://doi.org/10.1073/pnas.0511090103.
19. Cai Y-m, Hutchin A, Craddock J, Walsh MA, Webb JS, Tews I. 2020. Differential impact on motility and biofilm dispersal of closely related phosphodiesterases in *Pseudomonas aeruginosa*. Sci Rep 10:6232. https://doi.org/10.1038/s41598-020-63008-5.
20. Anantharaman V, Aravind L. 2000. Cache - A signaling domain common to animal Ca^{2+}-channel subunits and a class of prokaryotic chemotaxis receptors. Trends Biochem Sci 25:535–537. https://doi.org/10.1016/S0968-0004(00)01672-8.
21. Zhang Z, Hendrickson WA. 2010. Structural characterization of the predominant family of histidine kinase sensor domains. J Mol Biol 400:335–353. https://doi.org/10.1016/j.jmb.2010.04.049.
22. Upadhyay AA, Fleetwood AD, Adebali O, Finn RD, Zhulin IB. 2016. Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in Prokaryotes. PLoS Comput Biol 12:e1004862. https://doi.org/10.1371/journal.pcbi.1004862.
23. Cho US, Bader MW, Amaya MF, Daley ME, Klevit RE, Miller SI, Xu W. 2006. Metal bridges between the PhoQ sensor domain and the membrane regulate transmembrane signaling. J Mol Biol 356:1193–1206. https://doi.org/10.1016/j.jmb.2005.12.032.
24. Cheung J, Hendrickson WA. 2008. Crystal structures of C4-dicarboxylate ligand complexes with sensor domains of histidine kinases DcuS and DctB. J Biol Chem 283:30256–30265. https://doi.org/10.1074/jbc.M805253200.
25. Möglich A, Moffat K. 2007. Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. J Mol Biol 373:112–126. https://doi.org/10.1016/j.jmb.2007.07.039.
26. Shah N, Gaupp R, Moriyama H, Eskridge KM, Moriyama EN, Somerville G. 2013. Reductive evolution and the loss of PDC/PAS domains from the genus *Staphylococcus*. BMC Genomics 14:524. https://doi.org/10.1186/1471-2164-14-524.
27. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A. 2010. The Pfam protein families databases. Nucleic Acids Res 38:D211–D222. https://doi.org/10.1093/nar/gkp985.
28. Yoon B-J. 2009. Hidden Markov Models and their applications in biological sequence analysis. Curr Genomics 10:402–415. https://doi.org/10.2174/138920209789177575.
29. Remmert M, Biegert A, Hauser A, Söding J. 2011. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9:173–175. https://doi.org/10.1038/nmeth.1818.
30. Schultz J, Milpetz F, Bork P, Ponting CP. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A 95:5857–5864. https://doi.org/10.1073/pnas.95.11.5857.
31. Letunic I, Doerks T, Bork P. 2015. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res 43:D257–260. https://doi.org/10.1093/nar/gku949.
32. Hefti MH, Françoijs K-J, De Vries SC, Dixon R, Vervoort J. 2004. The PAS fold. Eur J Biochem 271:1198–1208. https://doi.org/10.1111/j.1432-1033.2004.04023.x.
33. Chen Y-T, Chang HY, Lu CL, Peng H-L. 2004. Evolutionary analysis of the two-component systems in *Pseudomonas aeruginosa* PAO1. J Mol Evol 59:725–737. https://doi.org/10.1007/s00239-004-2663-2.
34. Valentini M, Storelli N, Lapouge K. 2011. Identification of C4-dicarboxylate transport systems in *Pseudomonas aeruginosa* PAO1. J Bacteriol 193:4307–4316. https://doi.org/10.1128/JB.05074-11.
35. Shi R, McDonald L, Cygler M, Ekiel I. 2014. Coiled-coil helix rotation selects repressing or activating state of transcriptional regulator DhaR. Structure 22:478–487. https://doi.org/10.1016/j.str.2013.11.012.
36. García-Fontana C, Corral Lugo A, Krell T. 2014. Specificity of the CheR2 methyltransferase in *Pseudomonas aeruginosa* is directed by a C-terminal pentapeptide in the McpB chemoreceptor. Sci Signal 7:ra34. https://doi.org/10.1126/scisignal.2004849.

37. Pei J, Kim BH, Grishin NV. 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res 36:2295–2300. https://doi.org/10.1093/nar/gkn072.
38. King-Scott J, Konarev PV, Panjikar S, Jordanova R, Svergun DI, Tucker PA. 2011. Structural characterization of the multidomain regulatory protein Rv1364c from *Mycobacterium tuberculosis*. Structure 19:56–69. https://doi.org/10.1016/j.str.2010.11.010.
39. Wu H, Li M, Guo H, Zhou H, Li B, Xu Q, Xu C, Yu F, He J. 2019. Crystal structure of the *Vibrio cholerae* VqmA-ligand-DNA complex provides insight into ligand-binding mechanisms relevant for drug design. J Biol Chem 294:2580–2592. https://doi.org/10.1074/jbc.RA118.006082.
40. Bardy SL, Maddock JR. 2005. Polar localization of a soluble methyl-accepting protein of *Pseudomonas aeruginosa*. J Bacteriol 187:7840–7844. https://doi.org/10.1128/JB.187.22.7840-7844.2005.
41. Lundgren BR, Villegas-Peñaranda LR, Harris JR, Mottern AM, Dunn DM, Boddy CN, Nomura CT. 2014. Genetic analysis of the assimilation of C5-dicarboxylic acids in *Pseudomonas aeruginosa* PAO1. J Bacteriol 196:2543–2551. https://doi.org/10.1128/JB.01615-14.
42. Tatke G, Kumari H, Silva-Herzog E, Ramirez L, Mathee K. 2015. Pseudomonas aeruginosa a MifS-MifR two-component system is specific for α-ketoglutarate utilization. PLoS One 10:e0129629. https://doi.org/10.1371/journal.pone.0129629.
43. Sonawane AM, Singh B, Röhm KH. 2006. The AauR-AauS two-component system regulates uptake and metabolism of acidic amino acids in *Pseudomonas putida*. Appl Environ Microbiol 72:6569–6577. https://doi.org/10.1128/AEM.00830-06.
44. Gavira JA, Gumerov VM, Rico-Jiménez M, Petukh M, Upadhyay AA, Ortega A, Matilla MA, Zhulin IB, Krell T. 2020. How bacterial chemoreceptors evolve novel ligand specificities. mBio 11:e03066-19. https://doi.org/10.1128/mBio.03066-19.
45. Goers Sweeney E, Henderson JN, Goers J, Wreden C, Hicks KG, Foster JK, Parthasarathy R, Remington SJ, Guillemin K. 2012. Structure and proposed mechanism for the pH-sensing *Helicobacter pylori* chemoreceptor TlpB. Structure 20:1177–1188. https://doi.org/10.1016/j.str.2012.04.021.
46. Brewster JL, McKellar JLO, Finn TJ, Newman J, Peat TS, Gerth ML. 2016. Structural basis for ligand recognition by a Cache chemosensory domain that mediates carboxylate sensing in *Pseudomonas syringae*. Sci Rep 6:35198. https://doi.org/10.1038/srep35198.
47. Pokkuluri PR, Dwulit-Smith J, Duke NE, Wilton R, Mack JC, Bearden J, Rakowski E, Babnigg G, Szurmant H, Joachimiak A, Schiffer M. 2013. Analysis of periplasmic sensor domains from *Anaeromyxobacter dehalogenans* 2CP-C: structure of one sensor domain from a histidine kinase and another from a chemotaxis protein. MicrobiologyOpen 2:766–777. https://doi.org/10.1002/mbo3.112.
48. Alvarez-Ortega C, Harwood CS. 2007. Identification of a malate chemoreceptor in *Pseudomonas aeruginosa* by screening for chemotaxis defects in an energy taxis-deficient mutant. Appl Environ Microbiol 73:7793–7795. https://doi.org/10.1128/AEM.01898-07.
49. Lacal J, Alfonso C, Liu X, Parales RE, Morel B, Conejero-Lara F, Rivas G, Duque E, Ramos JL, Krell T. 2010. Identification of a chemoreceptor for tricarboxylic acid cycle intermediates: differential chemotactic response towards receptor ligands. J Biol Chem 285:23126–23136. https://doi.org/10.1074/jbc.M110.110403.
50. Reinelt S, Hofmann E, Gerharz T, Bott M, Madden DR. 2003. The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular PAS domain. J Biol Chem 278:39189–39196. https://doi.org/10.1074/jbc.M305864200.
51. Lundgren BR, Sarwar Z, Pinto A, Ganley JG, Nomura CT. 2016. Ethanolamine catabolism in *Pseudomonas aeruginosa* PAO1 is regulated by the enhancer-binding protein EatR (PA4021) and the alternative sigma factor RpoN. J Bacteriol 198:2318–2329. https://doi.org/10.1128/JB.00357-16.
52. Liu Q, Liu Y, Kang Z, Xiao D, Gao C, Xu P, Ma C. 2018. 2,3-Butanediol catabolism in *Pseudomonas aeruginosa* PAO1. Environ Microbiol 20:3927–3940. https://doi.org/10.1111/1462-2920.14332.
53. Camus L, Briaud P, Bastien S, Elsen S, Doléans-Jordheim A, Vandenesch F, Moreau K. 2020. Trophic cooperation promotes bacterial survival of *Staphylococcus aureus* and *Pseudomonas aeruginosa*. ISME J 14:3093–3105. https://doi.org/10.1038/s41396-020-00741-9.
54. Holm L, Elofsson A. 2019. Benchmarking fold detection by DaliLite v.5. Bioinformatics 35:5326–5327. https://doi.org/10.1093/bioinformatics/btz536.
55. Key J, Moffat K. 2005. Crystal structures of deoxy and CO-bound bjFixLH reveal details of ligand recognition and signaling. Biochemistry 44:4627–4635. https://doi.org/10.1021/bi047942r.
56. Rinaldi J, Gallo M, Klinke S, Paris G, Bonomi HR, Bogomolni RA, Cicero DO, Goldbaum FA. 2012. The β-scaffold of the LOV domain of the *Brucella* light-activated histidine kinase is a key element for signal transduction. J Mol Biol 420:112–127. https://doi.org/10.1016/j.jmb.2012.04.006.
57. Key J, Hefti M, Purcell EB, Moffat K. 2007. Structure of the redox sensor domain of *Azotobacter vinelandii* NifL at atomic resolution: signaling, dimerization, and mechanism. Biochemistry 46:3614–3623. https://doi.org/10.1021/bi0620407.
58. Borgstahl GE, Williams DR, Getzoff ED. 1995. 1.4 A structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site, and chromophore. Biochemistry 34:6278–6287. https://doi.org/10.1021/bi00019a004.
59. Cheung J, Bingman CA, Reyngold M, Hendrickson WA, Waldburger CD. 2008. Crystal structure of a functional dimer of the PhoQ sensor domain. J Biol Chem 283:13762–13770. https://doi.org/10.1074/jbc.M710592200.
60. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680. https://doi.org/10.1093/nar/22.22.4673.
61. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.
62. Kumar S, Tamura K, Nei M. 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Comput Appl Biosci 10:189–191. https://doi.org/10.1093/bioinformatics/10.2.189.
63. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454.
64. Gascuel O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 14:685–695. https://doi.org/10.1093/oxfordjournals.molbev.a025808.
65. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282. https://doi.org/10.1093/bioinformatics/8.3.275.
66. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.
67. Tian W, Chen C, Lei X, Zhao J, Liang J. 2018. CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res 46:W363–W367. https://doi.org/10.1093/nar/gky473.
68. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084.
69. Connolly ML. 1983. Solvent-accessible surfaces of proteins and nucleic acids. Science 221:709–713. https://doi.org/10.1126/science.6879170.
70. Rajagopal S, Moffat K. 2003. Crystal structure of a photoactive yellow protein from a sensor histidine kinase: conformational variability and signal transduction. Proc Natl Acad Sci U S A 100:1649–1654. https://doi.org/10.1073/pnas.0336353100.
71. Koh S, Hwang J, Guchhait K, Lee E-G, Kim S-Y, Kim S, Lee S, Chung JM, Jung HS, Lee SJ, Ryu C-M, Lee S-G, Oh T-K, Kwon O, Kim MH. 2016. Molecular insights into toluene sensing in the TodS/TodT signal transduction system. J Biol Chem 291:8575–8590. https://doi.org/10.1074/jbc.M116.718841.
72. Ukaegbu UE, Rosenzweig AC. 2009. Structure of the redox sensor domain of *Methylococcus capsulatus* (Bath) MmoS. Biochemistry 48:2207–2215. https://doi.org/10.1021/bi8019614.
73. Zoltowski BD, Schwerdtfeger C, Widom J, Loros JJ, Bilwes AM, Dunlap JC, Crane BR. 2007. Conformational switching in fungal light sensor Vivid. Science 316:1054–1057. https://doi.org/10.1126/science.1137128.
74. Xu Q, van Wezel GP, Chiu H-J, Jaroszewski L, Klock HE, Knuth MW, Miller MD, Lesley SA, Godzik A, Elsliger M-A, Deacon AM, Wilson IA. 2012. Structure of an MmyB-like regulator from *C. aurantiacus*, member of a new transcription factor family linked to antibiotic metabolism in actinomycetes. PLoS One 7:e41359. https://doi.org/10.1371/journal.pone.0041359.
75. Fala AM, Oliveira JF, Adamoski D, Aricetti JA, Dias MM, Dias MVB, Sforça ML, Lopes-de-Oliveira PS, Rocco SA, Caldana C, Dias SMG, Ambrosio ALB. 2015. Unsaturated fatty acids as high-affinity ligands of the C-terminal Per-ARNT-Sim domain from the Hypoxia-inducible factor 3α. Sci Rep 5:12698. https://doi.org/10.1038/srep12698.
76. Waldron EJ, Snyder D, Fernandez NL, Sileo E, Inoyama D, Freundlich JS, Waters CM, Cooper VS, Neiditch MB. 2019. Structural basis of DSF recognition by its receptor RpfR and its regulatory interaction with the DSF

synthase RpfF. PLoS Biol 17:e3000123. https://doi.org/10.1371/journal.pbio.3000123.
77. Banerjee A, Herman E, Kottke T, Essen LO. 2016. Structure of a native-like aureochrome 1a LOV domain dimer from *Phaeodactylum tricornutum*. Structure 24:171–178. https://doi.org/10.1016/j.str.2015.10.022.
78. Nazarenko VV, Remeeva A, Yudenko A, Kovalev K, Dubenko A, Goncharov IM, Kuzmichev P, Rogachev AV, Buslaev P, Borshchevskiy V, Mishin A, Dhoke GV, Schwaneberg U, Davari MD, Jaeger K-E, Krauss U, Gordeliy V, Gushchin I. 2019. A thermostable flavin-based fluorescent protein from *Chloroflexus aggregans*: a framework for ultra-high resolution structural studies. Photochem Photobiol Sci 18:1793–1805. https://doi.org/10.1039/c9pp00067d.
79. Fettweiss T, Röllen K, Granzin J, Reiners O, Endres S, Drepper T, Willbold D, Jaeger K-E, Batra-Safferling R, Krauss U. 2018. Mechanistic basis of the fast dark recovery of the short LOV protein DsLOV from *Dinoroseobacter shibae*. Biochemistry 57:4833–4847. https://doi.org/10.1021/acs.biochem.8b00645.
80. Nash AI, McNulty R, Shillito ME, Swartz TE, Bogomolni RA, Luecke H, Gardner KH. 2011. Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein. Proc Natl Acad Sci U S A 108:9449–9454. https://doi.org/10.1073/pnas.1100262108.
81. Rivera-Cancel G, Ko W-h, Tomchick DR, Correa F, Gardner KH. 2014. Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. Proc Natl Acad Sci U S A 111:17839–17844. https://doi.org/10.1073/pnas.1413983111.
82. Lokhandwala J, Hopkins HC, Rodriguez-Iglesias A, Dattenböck C, Schmoll M, Zoltowski BD. 2015. Structural biochemistry of a fungal LOV domain photoreceptor reveals an evolutionarily conserved pathway integrating light and oxidative stress. Structure 23:116–125. https://doi.org/10.1016/j.str.2014.10.020.
83. Conrad KS, Bilwes AM, Crane BR. 2013. Light-induced subunit dissociation by a light-oxygen-voltage domain photoreceptor from *Rhodobacter sphaeroides*. Biochemistry 52:378–391. https://doi.org/10.1021/bi3015373.
84. Halavaty AS, Moffat K. 2007. N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from *Avena sativa*. Biochemistry 46:14001–14009. https://doi.org/10.1021/bi701543e.
85. Kalvaitis ME, Johnson LA, Mart RJ, Rizkallah P, Allemann RK. 2019. A noncanonical chromophore reveals structural rearrangements of the Light-Oxygen-Voltage domain upon photoactivation. Biochemistry 58:2608–2616. https://doi.org/10.1021/acs.biochem.9b00255.
86. Weber AM, Kaiser J, Ziegler T, Pilsl S, Renzl C, Sixt L, Pietruschka G, Moniot S, Kakoti A, Juraschitz M, Schrottke S, Lledo Bryant L, Steegborn C, Bittl R, Mayer G, Möglich A. 2019. A blue light receptor that mediates RNA binding and translational regulation. Nat Chem Biol 15:1085–1092. https://doi.org/10.1038/s41589-019-0346-y.
87. Fedorov R, Schlichting I, Hartmann E, Domratcheva T, Fuhrmann M, Hegemann P. 2003. Crystal structures and molecular mechanism of a light-induced signaling switch: the Phot-LOV1 domain from *Chlamydomonas reinhardtii*. Biophys J 84:2474–2482. https://doi.org/10.1016/S0006-3495(03)75052-8.
88. Nakasako M, Zikihara K, Matsuoka D, Katsura H, Tokutomi S. 2008. Structural basis of the LOV1 dimerization of *Arabidopsis* phototropins 1 and 2. J Mol Biol 381:718–733. https://doi.org/10.1016/j.jmb.2008.06.033.
89. Crosson S, Moffat K. 2001. Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. Proc Natl Acad Sci U S A 98:2995–3000. https://doi.org/10.1073/pnas.051520298.
90. Circolone F, Granzin J, Jentzsch K, Drepper T, Jaeger K-E, Willbold D, Krauss U, Batra-Safferling R. 2012. Structural basis for the slow dark recovery of a full-length LOV protein from *Pseudomonas putida*. J Mol Biol 417:362–374. https://doi.org/10.1016/j.jmb.2012.01.056.
91. Mitra D, Yang X, Moffat K. 2012. Crystal structures of Aureochrome1 LOV suggest new design strategies for optogenetics. Structure 20:698–706. https://doi.org/10.1016/j.str.2012.02.016.
92. Pudasaini A, Shim JS, Song YH, Shi H, Kiba T, Somers DE, Imaizumi T, Zoltowski BD. 2017. Kinetics of the LOV domain of ZEITLUPE determine its circadian function in Arabidopsis. Elife 6:e21646. https://doi.org/10.7554/eLife.21646.
93. Sawai H, Sugimoto H, Shiro Y, Ishikawa H, Mizutani Y, Aono S. 2012. Structural basis for oxygen sensing and signal transduction of the heme-based sensor protein Aer2 from *Pseudomonas aeruginosa*. Chem Commun (Camb) 48:6523-5. https://doi.org/10.1039/c2cc32549g.
94. Greer-Phillips SE, Sukomon N, Chua TK, Johnson MS, Crane BR, Watts KJ. 2018. The Aer2 receptor from *Vibrio cholerae* is a dual PAS-heme oxygen sensor. Mol Microbiol 109:209–224. https://doi.org/10.1111/mmi.13978.
95. Kurokawa H, Lee D-S, Watanabe M, Sagami I, Mikami B, Raman CS, Shimizu T. 2004. A redox-controlled molecular switch revealed by the crystal structure of a bacterial heme PAS sensor. J Biol Chem 279:20186–20193. https://doi.org/10.1074/jbc.M314199200.
96. Ortmayer M, Lafite P, Menon BRK, Tralau T, Fisher K, Denkhaus L, Scrutton NS, Rigby SEJ, Munro AW, Hay S, Leys D. 2016. An oxidative N-demethylase reveals PAS transition from ubiquitous sensor to enzyme. Nature 539:593–597. https://doi.org/10.1038/nature20159.
97. Gong W, Hao B, Mansy SS, Gonzalez G, Gilles-Gonzalez MA, Chan MK. 1998. Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. Proc Natl Acad Sci U S A 95:15177–15182. https://doi.org/10.1073/pnas.95.26.15177.
98. Miyatake H, Mukai M, Park SY, Adachi S, Tamura K, Nakamura H, Nakamura K, Tsuchiya T, Iizuka T, Shiro Y. 2000. Sensory mechanism of oxygen sensor FixL from *Rhizobium meliloti*: crystallographic, mutagenesis and resonance Raman spectroscopic studies. J Mol Biol 301:415–431. https://doi.org/10.1006/jmbi.2000.3954.
99. Pokkuluri PR, Pessanha M, Londer YY, Wood SJ, Duke NEC, Wilton R, Catarino T, Salgueiro CA, Schiffer M. 2008. Structures and solution properties of two novel periplasmic sensor domains with c-type heme from chemotaxis proteins of *Geobacter sulfurreducens*: implications for signal transduction. J Mol Biol 377:1498–1517. https://doi.org/10.1016/j.jmb.2008.01.087.
100. Motomura T, Suga M, Hienerwadel R, Nakagawa A, Lai TL, Nitschke W, Kuma T, Sugiura M, Boussac A, Shen J-R. 2017. Crystal structure and redox properties of a novel cyanobacterial heme protein with a His/Cys heme axial ligation and a Per-Arnt-Sim (PAS)-like domain. J Biol Chem 292:9599–9612. https://doi.org/10.1074/jbc.M116.746263.
101. Affandi T, Issaian AV, McEvoy MM. 2016. The structure of the periplasmic sensor domain of the histidine kinase CusS shows unusual metal ion coordination at the dimeric interface. Biochemistry 55:5296–5306. https://doi.org/10.1021/acs.biochem.6b00707.
102. Wang D, Chen W, Huang S, He Y, Liu X, Hu Q, Wei T, Sang H, Gan J, Chen H. 2017. Structural basis of Zn(II) induced metal detoxification and antibiotic resistance by histidine kinase CzcS in *Pseudomonas aeruginosa*. PLoS Pathog 13:e1006533. https://doi.org/10.1371/journal.ppat.1006533.
103. Nagano S, Scheerer P, Zubow K, Michael N, Inomata K, Lamparter T, Krauß N. 2016. The crystal structures of the N-terminal photosensory core module of *Agrobacterium* phytochrome Agp1 as parallel and anti-parallel dimers. J Biol Chem 291:20674–20691. https://doi.org/10.1074/jbc.M116.739136.
104. Schmidt A, Sauthof L, Szczepek M, Lopez MF, Escobar FV, Qureshi BM, Michael N, Buhrke D, Stevens T, Kwiatkowski D, von Stetten D, Mroginski MA, Krauß N, Lamparter T, Hildebrandt P, Scheerer P. 2018. Structural snapshot of a bacterial phytochrome in its functional intermediate state. Nat Commun 9:4912. https://doi.org/10.1038/s41467-018-07392-7.
105. Schulte KW, Green E, Wilz A, Platten M, Daumke O. 2017. Structural basis for aryl hydrocarbon receptor-mediated gene activation. Structure 25:1025–1033.e3. https://doi.org/10.1016/j.str.2017.05.008.
106. Wu D, Potluri N, Kim Y, Rastinejad F. 2013. Structure and dimerization properties of the aryl hydrocarbon receptor PAS-A domain. Mol Cell Biol 33:4346–4356. https://doi.org/10.1128/MCB.00698-13.
107. Sakurai S, Shimizu T, Ohto U. 2017. The crystal structure of the AhRR-ARNT heterodimer reveals the structural basis of the repression of AhR-mediated transcription. J Biol Chem 292:17609–17616. https://doi.org/10.1074/jbc.M117.812974.
108. Card PB, Erbel PJA, Gardner KH. 2005. Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization. J Mol Biol 353:664–677. https://doi.org/10.1016/j.jmb.2005.08.043.
109. Wu D, Potluri N, Lu J, Kim Y, Rastinejad F. 2015. Structural integration in hypoxia-inducible factors. Nature 524:303–308. https://doi.org/10.1038/nature14883.
110. Huang N, Chelliah Y, Shan Y, Taylor CA, Yoo S-H, Partch C, Green CB, Zhang H, Takahashi JS. 2012. Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. Science 337:189–194. https://doi.org/10.1126/science.1222804.
111. Essen L-O, Mailliet J, Hughes J. 2008. The structure of a complete phytochrome sensory module in the Pr ground state. Proc Natl Acad Sci U S A 105:14709–14714. https://doi.org/10.1073/pnas.0806477105.

112. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD. 2005. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. Nature 438:325–331. https://doi.org/10.1038/nature04118.
113. Haitin Y, Carlson AE, Zagotta WN. 2013. The structural mechanism of KCNH-channel regulation by the eag domain. Nature 501:444–448. https://doi.org/10.1038/nature12487.
114. Whicher JR, MacKinnon R. 2016. Structure of the voltage-gated K^+ channel Eag1 reveals an alternative voltage sensing mechanism. Science 353:664–669. https://doi.org/10.1126/science.aaf8070.
115. Gourinchas G, Etzl S, Göbl C, Vide U, Madl T, Winkler A. 2017. Long-range allosteric signaling in red light-regulated diguanylyl cyclases. Sci Adv 3:e1602498. https://doi.org/10.1126/sciadv.1602498.
116. Russo L, Giller K, Pfitzner E, Griesinger C, Becker S. 2017. Insight into the molecular recognition mechanism of the coactivator NCoA1 by STAT6. Sci Rep 7:16845. https://doi.org/10.1038/s41598-017-17088-5.
117. Razeto A, Ramakrishnan V, Litterst CM, Giller K, Griesinger C, Carlomagno T, Lakomek N, Heimburg T, Lodrini M, Pfitzner E, Becker S. 2004. Structure of the NCoA-1/SRC-1 PAS-B domain bound to the LXXLL motif of the STAT6 transactivation domain. J Mol Biol 336:319–329. https://doi.org/10.1016/j.jmb.2003.12.057.
118. Yang X, Kuk J, Moffat K. 2008. Crystal structure of *Pseudomonas aeruginosa* bacteriophytochrome: photoconversion and signal transduction. Proc Natl Acad Sci U S A 105:14715–14720. https://doi.org/10.1073/pnas.0806718105.
119. Burgie ES, Bussell AN, Walker JM, Dubiel K, Vierstra RD. 2014. Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. Proc Natl Acad Sci U S A 111:10179–10184. https://doi.org/10.1073/pnas.1403096111.
120. Winkler A, Heintz U, Lindner R, Reinstein J, Shoeman RL, Schlichting I. 2013. A ternary AppA – PpsR – DNA complex mediates light-regulation of photosynthesis-related gene expression. Nat Struct Mol Biol 20:859–867. https://doi.org/10.1038/nsmb.2597.
121. Bellini D, Papiz MZ. 2012. Structure of a bacteriophytochrome and light-stimulated protomer swapping with a gene repressor. Structure 20:1436–1446. https://doi.org/10.1016/j.str.2012.06.002.
122. Bellini D, Papiz MZ. 2012. Dimerization properties of the RpBphP2 chromophore-binding domain crystallized by homologue-directed mutagenesis. Acta Crystallogr D Biol Crystallogr 68:1058–1066. https://doi.org/10.1107/S0907444912020537.
123. Yang X, Stojkovic EA, Kuk J, Moffat K. 2007. Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. Proc Natl Acad Sci U S A 104:12571–12576. https://doi.org/10.1073/pnas.0701737104.
124. Woitowich NC, Halavaty AS, Waltz P, Kupitz C, Valera J, Tracy G, Gallagher KD, Claesson E, Nakane T, Pandey S, Nelson G, Tanaka R, Nango E, Mizohata E, Owada S, Tono K, Joti Y, Nugent AC, Patel H, Mapara A, Hopkins J, Duong P, Bizhga D, Kovaleva SE, St Peter R, Hernandez CN, Ozarowski WB, Roy-Chowdhuri S, Yang J-H, Edlund P, Takala H, Ihalainen J, Brayshaw J, Norwood T, Poudyal I, Fromme P, Spence JCH, Moffat K, Westenhoff S, Schmidt M, Stojković EA. 2018. Structural basis for light control of cell development revealed by crystal structures of a myxobacterial phytochrome. IUCrJ 5:619–634. https://doi.org/10.1107/S2052252518010631.
125. Sanchez JC, Carrillo M, Pandey S, Noda M, Aldama L, Feliz D, Claesson E, Wahlgren WY, Tracy G, Duong P, Nugent AC, Field A, Šrajer V, Kupitz C, Iwata S, Nango E, Tanaka R, Tanaka T, Fangjia L, Tono K, Owada S, Westenhoff S, Schmidt M, Stojković EA. 2019. High-resolution crystal structures of a myxobacterial phytochrome at cryo and room temperatures. Struct Dyn 6:054701. https://doi.org/10.1063/1.5120527.
126. Purohit R, Weichsel A, Montfort WR. 2013. Crystal structure of the Alpha subunit PAS domain from soluble guanylyl cyclase. Protein Sci 22:1439–1444. https://doi.org/10.1002/pro.2331.
127. Kang Y, Liu R, Wu J-X, Chen L. 2019. Structural insights into the mechanism of human soluble guanylate cyclase. Nature 574:206–210. https://doi.org/10.1038/s41586-019-1584-6.
128. Otero LH, Klinke S, Rinaldi J, Velázquez-Escobar F, Mroginski MA, Fernández López M, Malamud F, Vojnov AA, Hildebrandt P, Goldbaum FA, Bonomi HR. 2016. Structure of the full-length bacteriophytochrome from the plant pathogen *Xanthomonas campestris* provides clues to its long-range signaling mechanism. J Mol Biol 428:3702–3720. https://doi.org/10.1016/j.jmb.2016.04.012.
129. Ud-Din AIMS, Khan MF, Roujeinikova A. 2020. Broad specificity of amino acid chemoreceptor CtaA of *Pseudomonas fluorescens* is afforded by plasticity of its amphipathic ligand-binding pocket. Mol Plant Microbe Interact 33:612–623. https://doi.org/10.1094/MPMI-10-19-0277-R.
130. Nishiyama S-i, Takahashi Y, Yamamoto K, Suzuki D, Itoh Y, Sumita K, Uchida Y, Homma M, Imada K, Kawagishi I. 2016. Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. Sci Rep 6:20866. https://doi.org/10.1038/srep20866.
131. Liu YC, Machuca MA, Beckham SA, Gunzburg MJ, Roujeinikova A. 2015. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. Acta Crystallogr D Biol Crystallogr 71:2127–2136. https://doi.org/10.1107/S139900471501384X.
132. Corral-Lugo A, Matilla MA, Martín-Mora D, Silva Jiménez H, Mesa Torres N, Kato J, Hida A, Oku S, Conejero-Muriel M, Gavira JA, Krell T. 2018. High-affinity chemotaxis to histamine mediated by the TlpQ chemoreceptor of the human pathogen *Pseudomonas aeruginosa*. mBio 9:e01894-18. https://doi.org/10.1128/mBio.01894-18.
133. Wu R, Wilton R, Cuff ME, Endres M, Babnigg G, Edirisinghe JN, Henry CS, Joachimiak A, Schiffer M, Pokkuluri PR. 2017. A novel signal transduction protein: combination of solute binding and tandem PAS-like sensor domains in one polypeptide chain. Protein Sci 26:857–869. https://doi.org/10.1002/pro.3134.
134. Gavira JA, Ortega Á, Martín-Mora D, Conejero-Muriel MT, Corral-Lugo A, Morel B, Matilla MA, Krell T. 2018. Structural basis for polyamine binding at the dCACHE domain of the McpU chemoreceptor from *Pseudomonas putida*. J Mol Biol 430:1950–1963. https://doi.org/10.1016/j.jmb.2018.05.008.
135. Neiditch MB, Federle MJ, Pompeani AJ, Kelly RC, Swem DL, Jeffrey PD, Bassler BL, Hughson FM. 2006. Ligand-induced asymmetry in histidine sensor kinase complex regulates quorum sensing. Cell 126:1095–1108. https://doi.org/10.1016/j.cell.2006.07.032.
136. Shrestha M, Compton KK, Mancl JM, Webb BA, Brown AM, Scharf BE, Schubot FD. 2018. Structure of the sensory domain of McpX from *Sinorhizobium meliloti*, the first known bacterial chemotactic sensor for quaternary ammonium compounds. Biochem J 475:3949–3962. https://doi.org/10.1042/BCJ20180769.
137. Hothorn M, Dabi T, Chory J. 2011. Structural basis for cytokinin recognition by *Arabidopsis thaliana* histidine kinase 4. Nat Chem Biol 7:766–768. https://doi.org/10.1038/nchembio.667.
138. Zhou YF, Nan B, Nan J, Ma Q, Panjikar S, Liang YH, Wang Y, Su XD. 2008. C4-Dicarboxylates sensing mechanism revealed by the crystal structures of DctB sensor domain. J Mol Biol 383:49–61. https://doi.org/10.1016/j.jmb.2008.08.010.
139. Wu R, Gu M, Wilton R, Babnigg G, Kim Y, Pokkuluri PR, Szurmant H, Joachimiak A, Schiffer M. 2013. Insight into the sporulation phosphorelay: crystal structure of the sensor domain of *Bacillus subtilis* histidine kinase, KinD. Protein Sci 22:564–576. https://doi.org/10.1002/pro.2237.
140. Machuca MA, Johnson KS, Liu YC, Steer DL, Ottemann KM, Roujeinikova A. 2017. Helicobacter pylori chemoreceptor TlpC mediates chemotaxis to lactate. Sci Rep 7:14089. https://doi.org/10.1038/s41598-017-14372-2.
141. Kim H-E, Shitashiro M, Kuroda A, Takiguchi N, Kato J. 2007. Ethylene chemotaxis in *Pseudomonas aeruginosa* and other *Pseudomonas* species. Microb Environ 22:186–189. https://doi.org/10.1264/jsme2.22.186.
142. Kato J, Kim H-E, Takiguchi N, Kuroda A, Ohtake H. 2008. *Pseudomonas aeruginosa* as a model microorganism for investigation of chemotactic behaviors in ecosystem. J Biosci Bioeng 106:1–7. https://doi.org/10.1263/jbb.106.1.
143. Taguchi K, Fukutomi H, Kuroda A, Kato J, Ohtake H. 1997. Genetic identification of chemotactic transducers for amino acids in *Pseudomonas aeruginosa*. Microbiology 143:3223–3229. https://doi.org/10.1099/00221287-143-10-3223.
144. Rico-Jiménez M, Muñoz-Martínez F, García-Fontana C, Fernandez M, Morel B, Ortega A, Ramos JL, Krell T. 2013. Paralogous chemoreceptors mediate chemotaxis towards protein amino acids and the non-protein amino acid gamma-aminobutyrate (GABA). Mol Microbiol 88:1230–1243. https://doi.org/10.1111/mmi.12255.
145. McKellar JLO, Minnell JJ, Gerth ML. 2015. A high-throughput screen for ligand binding reveals the specificities of three amino acid chemoreceptors from *Pseudomonas syringae* pv. *actinidiae*. Mol Microbiol 96:694–707. https://doi.org/10.1111/mmi.12964.
146. Dasgupta N, Wolfgang MC, Goodman AL, Arora SK, Jyot J, Lory S, Ramphal R. 2003. A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa*. Mol Microbiol 50:809–824. https://doi.org/10.1046/j.1365-2958.2003.03740.x.

147. Gallagher LA, McKnight SL, Kuznetsova MS, Pesci EC, Manoil C. 2002. Functions required for extracellular quinolone signaling by *Pseudomonas aeruginosa*. J Bacteriol 184:6472–6480. https://doi.org/10.1128/JB.184.23.6472-6480.2002.
148. Francis VI, Stevenson EC, Porter SL. 2017. Two-component systems required for virulence in *Pseudomonas aeruginosa*. FEMS Microbiology Lett 364. https://doi.org/10.1093/femsle/fnx104.
149. Li G, Lu C-D. 2016. Molecular characterization of LhpR in control of hydroxyproline catabolism and transport in *Pseudomonas aeruginosa* PAO1. Microbiology (Reading) 162:1232–1242. https://doi.org/10.1099/mic.0.000300.
150. Lundgren BR, Harris JR, Sarwar Z, Scheel RA, Nomura CT. 2015. The metabolism of (*R*)-3-hydroxybutyrate is regulated by the enhancer-binding protein PA2005 and the alternative sigma factor RpoN in *Pseudomonas aeruginosa* PAO1. Microbiology (Reading) 161:2232–2242. https://doi.org/10.1099/mic.0.000163.
151. Martín-Mora D, Ortega Á, Pérez-Maldonado FJ, Krell T, Matilla MA. 2018. The activity of the C4-dicarboxylic acid chemoreceptor of *Pseudomonas aeruginosa* is controlled by chemoattractants and antagonists. Sci Rep 8:2102. https://doi.org/10.1038/s41598-018-20283-7.
152. Mern DS, Ha SW, Khodaverdi V, Gliese N, Görisch H. 2010. A complex regulatory network controls aerobic ethanol oxidation in *Pseudomonas aeruginosa*: indication of four levels of sensor kinases and response regulators. Microbiology (Reading) 156:1505–1516. https://doi.org/10.1099/mic.0.032847-0.
153. Beaudoin T, Zhang L, Hinz AJ, Parr CJ, Mah T-F. 2012. The biofilm-specific antibiotic resistance gene ndvB is important for expression of ethanol oxidation genes in *Pseudomonas aeruginosa* biofilms. J Bacteriol 194:3128–3136. https://doi.org/10.1128/JB.06178-11.
154. Wei Q, Leclercq S, Bhasme P, Xu A, Zhu B, Zhang Y, Zhang M, Wang S, Ma LZ. 2019. Diguanylate cyclases and phosphodiesterases required for basal-level c-di-GMP in *Pseudomonas aeruginosa* as revealed by systematic phylogenetic and transcriptomic analyses. Appl Environ Microbiol 85:e01194-19. https://doi.org/10.1128/AEM.01194-19.
155. Barraud N, Schleheck D, Klebensberger J, Webb JS, Hassett DJ, Rice S, Kjelleberg S. 2009. Nitric oxide signaling in *Pseudomonas aeruginosa* biofilms mediates phosphodiesterase activity, decreased cyclic di-GMP levels, and enhanced dispersal. J Bacteriol 191:7333–7342. https://doi.org/10.1128/JB.00975-09.
156. Li Y, Petrova OE, Su S, Lau GW, Panmanee W, Na R, Hassett DJ, Davies DG, Sauer K. 2014. BdlA, DipA and induced dispersion contribute to acute virulence and chronic persistence of *Pseudomonas aeruginosa*. PLoS Pathog 10:e1004168. https://doi.org/10.1371/journal.ppat.1004168.
157. Hong CS, Kuroda A, Ikeda T, Takiguchi N, Ohtake H, Kato J. 2004. The aerotaxis transducer gene aer, but not aer-2, is transcriptionally regulated by the anaerobic regulator ANR in *Pseudomonas aeruginosa*. J Biosci Bioeng 97:184–190. https://doi.org/10.1016/S1389-1723(04)70188-7.
158. Hong CS, Shitashiro M, Kuroda A, Ikeda T, Takiguchi N, Ohtake H, Kato J. 2004. Chemotaxis proteins and transducers for aerotaxis in *Pseudomonas aeruginosa*. FEMS Microbiol Lett 231:247–252. https://doi.org/10.1016/S0378-1097(04)00009-6.
159. Choy W-K, Zhou L, Syn CK-C, Zhang L-H, Swarup S. 2004. MorA defines a new class of regulators affecting flagellar development and biofilm formation in diverse *Pseudomonas* species. J Bacteriol 186:7221–7228. https://doi.org/10.1128/JB.186.21.7221-7228.2004.
160. Meissner A, Wild V, Simm R, Rohde M, Erck C, Bredenbruch F, Morr M, Römling U, Häussler S. 2007. *Pseudomonas aeruginosa* cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate. Environ Microbiol 9:2475–2485. https://doi.org/10.1111/j.1462-2920.2007.01366.x.
161. Phippen CW, Mikolajek H, Schlaefli HG, Keevil CW, Webb JS, Tews I. 2014. Formation and dimerization of the phosphodiesterase active site of the *Pseudomonas aeruginosa* MorA, a bi-functional c-di-GMP regulator. FEBS Lett 588:4631–4636. https://doi.org/10.1016/j.febslet.2014.11.002.
162. Ravichandran A, Ramachandran M, Suriyanarayanan T, Wong CC, Swarup S. 2015. Global regulator MorA affects virulence-associated protease secretion in *Pseudomonas aeruginosa* PAO1. PLoS One 10:e0123805. https://doi.org/10.1371/journal.pone.0123805.
163. Waite RD, Papakonstantinopoulou A, Littler E, Curtis MA. 2005. Transcriptome analysis of *Pseudomonas aeruginosa* growth: comparison of gene expression in planktonic cultures and developing and mature biofilms. J Bacteriol 187:6571–6576. https://doi.org/10.1128/JB.187.18.6571-6576.2005.
164. Müsken M, Di Fiore S, Dötsch A, Fischer R, Häussler S. 2010. Genetic determinants of *Pseudomonas aeruginosa* biofilm establishment. Microbiology (Reading) 156:431–441. https://doi.org/10.1099/mic.0.033290-0.
165. Bhasme P, Wei Q, Xu A, Naqvi STA, Wang D, Ma LZ. 2020. Evaluation and characterization of the predicted diguanylate cyclase-encoding genes in *Pseudomonas aeruginosa*. MicrobiologyOpen 9:e975. https://doi.org/10.1002/mbo3.975.
166. Zhang Y, Guo J, Zhang N, Yuan W, Lin Z, Huang W. 2019. Characterization and analysis of a novel diguanylate cyclase PA0847 from *Pseudomonas aeruginosa* PAO1. Infect Drug Resist 12:655–665. https://doi.org/10.2147/IDR.S194462.
167. Liang H, Li L, Kong W, Shen L, Duan K. 2009. Identification of a novel regulator of the quorum-sensing systems in *Pseudomonas aeruginosa*. FEMS Microbiol Lett 293:196–204. https://doi.org/10.1111/j.1574-6968.2009.01544.x.
168. Lundgren BR, Bailey FJ, Moley G, Nomura CT. 2017. DdaR (PA1196) regulates expression of dimethylarginine dimethylaminohydrolase for the metabolism of methylarginines in *Pseudomonas aeruginosa* PAO1. J Bacteriol 199:e00001-17. https://doi.org/10.1128/JB.00001-17.
169. Badal D, Jayarani AV, Kollaran MA, Kumar A, Singh V. 2020. *Pseudomonas aeruginosa* biofilm formation on endotracheal tubes requires multiple two-component systems. J Med Microbiol 69:906–919. https://doi.org/10.1099/jmm.0.001199.
170. Ranjitkar S, Jones AK, Mostafavi M, Zwirko Z, Iartchouk O, Whitney Barnes S, Walker JR, Willis TW, Lee PS, Dean CR. 2019. Target (MexB)- and efflux-based mechanisms decreasing the effectiveness of the efflux pump inhibitor D13-9001 in Pseudomonas aeruginosa PAO1: uncovering a new role for MexMn-OpRM in efflux of β-lactams and a novel regulatory circuit (MmnRS) controlling Me. Antimicrob Agents Chemother 63:e01718-18. https://doi.org/10.1128/AAC.01718-18.
171. Petrova OE, Sauer K. 2009. A novel signaling network essential for regulating *Pseudomonas aeruginosa* biofilm development. PLoS Pathog 5:e1000668. https://doi.org/10.1371/journal.ppat.1000668.
172. Petrova OE, Sauer K. 2010. The novel two-component regulatory system BfiSR regulates biofilm development by controlling the small RNA *rsmZ* through CafA. J Bacteriol 192:5275–5288. https://doi.org/10.1128/JB.00387-10.
173. Petrova OE, Sauer K. 2011. SagS contributes to the motile-sessile switch and acts in concert with BfiSR to enable *Pseudomonas aeruginosa* biofilm formation. J Bacteriol 193:6614–6628. https://doi.org/10.1128/JB.00305-11.
174. Petrova OE, Gupta K, Liao J, Goodwine JS, Sauer K. 2017. Divide and conquer: the *Pseudomonas aeruginosa* two-component hybrid SagS enables biofilm formation and recalcitrance of biofilm cells to antimicrobial agents via distinct regulatory circuits. Environ Microbiol 19:2005–2024. https://doi.org/10.1111/1462-2920.13719.
175. Wang Y, Ha U, Zeng L, Jin S. 2003. Regulation of membrane permeability by a two-component regulatory system in *Pseudomonas aeruginosa*. Antimicrob Agents Chemother 47:95–101. https://doi.org/10.1128/AAC.47.1.95-101.2003.
176. Giraud C, Bernard CS, Calderon V, Yang L, Filloux A, Molin S, Fichant G, Bordi C, de Bentzmann S. 2011. The PprA–PprB two-component system activates CupE, the first non-archetypal *Pseudomonas aeruginosa* chaperone–usher pathway system assembling fimbriae. Environ Microbiol 13:666–683. https://doi.org/10.1111/j.1462-2920.2010.02372.x.
177. de Bentzmann S, Giraud C, Bernard CS, Calderon V, Ewald F, Plésiat P, Nguyen C, Grunwald D, Attree I, Jeannot K, Fauvarque M-O, Bordi C. 2012. Unique biofilm signature, drug susceptibility and decreased virulence in *Drosophila* through the *Pseudomonas aeruginosa* two-component system PprAB. PLoS Pathog 8:e1003052. https://doi.org/10.1371/journal.ppat.1003052.
178. Genschik P, Drabikowski K, Filipowicz W. 1998. Characterization of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its sigma54-regulated operon. J Biol Chem 273:25516–25526. https://doi.org/10.1074/jbc.273.39.25516.
179. Winsor GL, Lam DKW, Fleming L, Lo R, Whiteside MD, Yu NY, Hancock REW, Brinkman FSL. 2011. Pseudomonas Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. Nucleic Acids Res 39:D596–600. https://doi.org/10.1093/nar/gkq869.
180. Klockgether J, Miethke N, Kubesch P, Bohn YS, Brockhausen I, Cramer N, Eberl L, Greipel J, Herrmann C, Herrmann S, Horatzek S, Lingner M, Luciano L, Salunkhe P, Schomburg D, Wehsling M, Wiehlmann L, Davenport CF, Tümmler B. 2013. Intraclonal diversity of the *Pseudomonas aeruginosa* cystic

fibrosis airway isolates TBCF10839 and TBCF121838: distinct signatures of transcriptome, proteome, metabolome, adherence and pathogenicity despite an almost identical genome sequence. Environ Microbiol 15:191–210. https://doi.org/10.1111/j.1462-2920.2012.02842.x.

181. Li Y, Xia H, Bai F, Xu H, Yang L, Yao H, Zhang L, Zhang X, Bai Y, Saris PEJ, Tolker-Nielsen T, Qiao M. 2007. Identification of a new gene PA5017 involved in flagella-mediated motility, chemotaxis and biofilm formation in *Pseudomonas aeruginosa*. FEMS Microbiol Lett 272:188–195. https://doi.org/10.1111/j.1574-6968.2007.00753.x.
182. Roy AB, Petrova OE, Sauer K. 2012. The phosphodiesterase DipA (PA5017) is essential for *Pseudomonas aeruginosa* biofilm dispersion. J Bacteriol 194:2904–2915. https://doi.org/10.1128/JB.05346-11.
183. Kulasekara BR, Kamischke C, Kulasekara HD, Christen M, Wiggins PA, Miller SI. 2013. c-di-GMP heterogeneity is generated by the chemotaxis machinery to regulate flagellar motility. Elife 2:e01402. https://doi.org/10.7554/eLife.01402.
184. Airola MV, Huh D, Sukomon N, Widom J, Sircar R, Borbat PP, Freed JH, Watts KJ, Crane BR. 2013. Architecture of the soluble receptor aer2 indicates an in-line mechanism for PAS and HAMP domain signaling. J Mol Biol 425:886–901. https://doi.org/10.1016/j.jmb.2012.12.011.
185. Luu RA, Schneider BJ, Ho CC, Nesteryuk V, Ngwesse SE, Liu X, Parales JV, Ditty JL, Parales RE. 2013. Taxis of *Pseudomonas putida* F1 toward phenylacetic acid is mediated by the energy taxis receptor AER2. Appl Environ Microbiol 79:2416–2423. https://doi.org/10.1128/AEM.03895-12.
186. Zaoui C, Overhage J, Löns D, Zimmermann A, Müsken M, Bielecki P, Pustelny C, Becker T, Nimtz M, Häussler S. 2012. An orphan sensor kinase controls quinolone signal production via MexT in *Pseudomonas aeruginosa*. Mol Microbiol 83:536–547. https://doi.org/10.1111/j.1365-2958.2011.07947.x.
187. Jacob K, Rasmussen A, Tyler P, Servos MM, Sylla M, Prado C, Daniele E, Sharp JS, Purdy AE. 2017. Regulation of acetyl-CoA synthetase transcription by the CrbS/R two-component system is conserved in genetically diverse environmental pathogens. PLoS One 12:e0177825. https://doi.org/10.1371/journal.pone.0177825.
188. Zamorano L, Moyà B, Juan C, Mulet X, Blázquez J, Oliver A. 2014. The *Pseudomonas aeruginosa* CreBC two-component system plays a major role in the response to β-lactams, fitness, biofilm growth, and global regulation. Antimicrob Agents Chemother 58:5084–5095. https://doi.org/10.1128/AAC.02556-14.
189. Godoy MS, Nikel PI, Cabrera Gomez JG, Julia Pettinari M. 2016. The CreC regulator of *Escherichia coli*, a new target for metabolic manipulations. Appl Environ Microbiol 82:244–254. https://doi.org/10.1128/AEM.02984-15.
190. Hassan MT, van der Lelie D, Springael D, Römling U, Ahmed N, Mergeay M. 1999. Identification of a gene cluster, *czr*, involved in cadmium and zinc resistance in *Pseudomonas aeruginosa*. Gene 238:417–425. https://doi.org/10.1016/s0378-1119(99)00349-2.
191. Perron K, Caille O, Rossier C, Van Delden C, Dumas JL, Köhler T. 2004. CzcR-CzcS, a two-component system involved in heavy metal and carbapenem resistance in *Pseudomonas aeruginosa*. J Biol Chem 279:8761–8768. https://doi.org/10.1074/jbc.M312080200.
192. Caille O, Rossier C, Perron K. 2007. A copper-activated two-component system interacts with zinc and imipenem resistance in *Pseudomonas aeruginosa*. J Bacteriol 189:4561–4568. https://doi.org/10.1128/JB.00095-07.
193. Bouffartigues E, Moscoso JA, Duchesne R, Rosay T, Fito-Boncompte L, Gicquel G, Maillot O, Bénard M, Bazire A, Brenner-Weiss G, Lesouhaitier O, Lerouge P, Dufour A, Orange N, Feuilloley MGJ, Overhage J, Filloux A, Chevalier S. 2015. The absence of the *Pseudomonas aeruginosa* OprF protein leads to increased biofilm formation through variation in c-di-GMP level. Front Microbiol 6:630. https://doi.org/10.3389/fmicb.2015.00630.
194. Tasler R, Moises T, Frankenberg-Dinkel N. 2005. Biochemical and spectroscopic characterization of the bacterial phytochrome of *Pseudomonas aeruginosa*. FEBS J 272:1927–1936. https://doi.org/10.1111/j.1742-4658.2005.04623.x.
195. Barkovits K, Schubert B, Heine S, Scheer M, Frankenberg-Dinkel N. 2011. Function of the bacteriophytochrome BphP in the RpoS/Las quorum-sensing network of *Pseudomonas aeruginosa*. Microbiology (Reading) 157:1651–1664. https://doi.org/10.1099/mic.0.049007-0.
196. Merrick MJ, Edwards RA. 1995. Nitrogen control in bacteria. Microbiol Rev 59:604–622. https://doi.org/10.1128/mr.59.4.604-622.1995.
197. Nishijyo T, Haas D, Itoh Y. 2001. The CbrA-CbrB two-component regulatory system controls the utilization of multiple carbon and nitrogen sources in *Pseudomonas aeruginosa*. Mol Microbiol 40:917–931. https://doi.org/10.1046/j.1365-2958.2001.02435.x.
198. Li W, Lu CD. 2007. Regulation of carbon and nitrogen utilization by CbrAB and NtrBC two-component systems in *Pseudomonas aeruginosa*. J Bacteriol 189:5413–5420. https://doi.org/10.1128/JB.00432-07.
199. Yeung ATY, Torfs ECW, Jamshidi F, Bains M, Wiegand I, Hancock REW, Overhage J. 2009. Swarming of *Pseudomonas aeruginosa* is controlled by a broad spectrum of transcriptional regulators, including MetR. J Bacteriol 191:5592–5602. https://doi.org/10.1128/JB.00157-09.
200. Alford MA, Baghela A, Yeung ATY, Pletzer D, Hancock REW. 2020. NtrBC regulates invasiveness and virulence of *Pseudomonas aeruginosa* during high-density infection. Front Microbiol 11:773. https://doi.org/10.3389/fmicb.2020.00773.