Multi-PAS domain-mediated protein oligomerization of PpsR from *Rhodobacter sphaeroides*

Udo Heintz, Anton Meinhart and Andreas Winkler*

Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Heidelberg, Germany

Correspondence e-mail: andreas.winkler@mpimf-heidelberg.mpg.de

Per–ARNT–Sim (PAS) domains are essential modules of many multi-domain signalling proteins that mediate protein interaction and/or sense environmental stimuli. Frequently, multiple PAS domains are present within single polypeptide chains, where their interplay is required for protein function. Although many isolated PAS domain structures have been reported over the last decades, only a few structures of multi-PAS proteins are known. Therefore, the molecular mechanism of multi-PAS domain-mediated protein oligomerization and function is poorly understood. The transcription factor PpsR from *Rhodobacter sphaeroides* is such a multi-PAS domain protein that, in addition to its three PAS domains, contains a glutamine-rich linker and a C-terminal helix–turn–helix DNA-binding motif. Here, crystal structures of two N-terminally and C-terminally truncated PpsR variants that comprise a single (PpsR$_{Q-PAS1}$) and two (PpsR$_{N-Q-PAS1}$) PAS domains, respectively, are presented and the multi-step strategy required for the phasing of a triple PAS domain construct (PpsR$_{\Delta HTH}$) is illustrated. While parts of the biologically relevant dimerization interface can already be observed in the two shorter constructs, the PpsR$_{\Delta HTH}$ structure reveals how three PAS domains enable the formation of multiple oligomeric states (dimer, tetramer and octamer), highlighting that not only the PAS cores but also their α-helical extensions are essential for protein oligomerization. The results demonstrate that the long helical glutamine-rich linker of PpsR results from a direct fusion of the N-cap of the PAS1 domain with the C-terminal extension of the N-domain that plays an important role in signal transduction.

1. Introduction

Protein self-association into homodimers or higher oligomeric states is essential for the function of many proteins and confers several advantages compared with monomers (Ali & Imperiali, 2005; Matthews & Sunde, 2012). Protein oligomers are not only more resistant to protein denaturation and proteolytic degradation, but also provide additional possibilities for functional control, such as allosteric regulation or oligomer-dependent activity. The oligomerization of many signalling proteins is mediated *via* Per–ARNT–Sim (PAS) domains. PAS domains are 100–120-amino-acid structural modules that occur in all kingdoms of life. Although PAS domains can exhibit low sequence identities among each other, they share a conserved three-dimensional fold consisting of a five-stranded antiparallel β-sheet surrounded by several short α-helices (Pellequer *et al.*, 1998; Hefti *et al.*, 2004). In addition to the structurally conserved PAS core, most PAS domains possess α-helical extensions that often connect multiple PAS domains within the same protein and/or link them to a variety of effector domains such as histidine kinases,

nucleotide cyclases, phosphodiesterases or DNA-binding domains (Möglich *et al.*, 2009*b*). This reflects their central involvement in the regulation of a variety of biological processes, such as the maintenance of circadian rhythms, control of phototropism, nitrogen fixation and gene transcription (Taylor & Zhulin, 1999; Möglich *et al.*, 2009*b*; Henry & Crosson, 2011). The α-helical PAS extensions can be found at the N-termini as well as the C-termini of PAS domains and commonly exhibit a heptad pattern of hydrophobic residues typical of amphipathic helices or coiled coils (Möglich *et al.*, 2009*b*). Owing to the importance of the N-terminal helices for the stabilization of PAS dimers in a parallel arrangement and its consequences for PAS function, these elements are generally referred to as 'N-caps' (Ayers & Moffat, 2008). Recently, it has been proposed that helical PAS extensions that connect different domains within a PAS protein are also involved in signal transmission in addition to protein oligomerization (Little *et al.*, 2012).

Some PAS domains bind cofactors and ligands, such as haem, flavins, carboxylic acids or divalent metal ions, in a cleft formed by the inner side of the β-sheet and the short helices (Henry & Crosson, 2011). These interactions provide the possibility of sensing a plethora of different signals, which result in conformational changes of the PAS domain. These signals are then further transmitted to downstream effector domains, modulating their activity. Alternatively, PAS proteins mediate signal transduction or regulate gene transcription by protein hetero- or homo-oligomerization as, for example, in histidine kinases (Krell *et al.*, 2010), light, oxygen, voltage (LOV) proteins (Herrou & Crosson, 2011) and basic helix–loop–helix/PAS transcription regulators (Kewley *et al.*, 2004). However, although many PAS proteins such as KinA, WC-1 (White Collar 1) and ARNT contain multiple PAS domains, the molecular mechanism of multi-PAS domain-mediated protein oligomerization is only poorly understood.

The redox-responding transcription factor PpsR (photopigment suppression) from *Rhodobacter sphaeroides* (Penfold & Pemberton, 1991) is such a multi-PAS domain-containing protein (Fig. 1). PpsR possesses three PAS domains (designated N-domain, PAS1 and PAS2) and a highly conserved glutamine-rich linker (Q-linker) as well as a C-terminal helix–turn–helix (HTH) DNA-binding motif (Fig. 2). PpsR acts as a repressor of photosynthesis and tetrapyrrole-biosynthesis gene expression by binding to palindromic DNA sequences (TGTc-N_{10}-gACA) present in the promoter region of PpsR-

regulated genes (Moskvin *et al.*, 2005; Elsen *et al.*, 2005). It has been proposed that redox-dependent DNA binding of PpsR is regulated by the formation of an intramolecular disulfide bond (Masuda & Bauer, 2002), as well as by the binding of a haem cofactor (Yin *et al.*, 2012). Additionally, the function of PpsR is regulated in a blue-light- and oxygen-dependent manner by the antirepressor AppA (Gomelsky & Kaplan, 1998; Braatsch *et al.*, 2002; Masuda & Bauer, 2002; Han *et al.*, 2007) that, together with PpsR and its cognate DNA, is able to form a light-sensitive ternary AppA–PpsR$_2$–DNA complex (Winkler *et al.*, 2013).

Here, we report the determination of the three-dimensional structure of PpsR from *R. sphaeroides*. Based on secondary-structure predictions, we designed N- and C-terminally truncated PpsR variants that comprise a single, two and three PAS domains (designated PpsR$_{Q-PAS1}$, PpsR$_{N-Q-PAS1}$ and PpsR$_{\Delta HTH}$ respectively; Fig. 1) to complement our crystallization approach of PpsR$_{full}$. We solved crystal structures of the truncated PpsR variants and describe the multi-step strategy required for phasing of PpsR$_{\Delta HTH}$. The structural studies are complemented by multi-angle light-scattering (MALS) and microscale thermophoresis (MST) measurements to determine and quantify the oligomerization states of the PpsR variants in solution. In conclusion, the PpsR structures reveal an intriguing PAS-mediated protein homo-oligomerization, demonstrating that the α-helical extensions significantly contribute to protein oligomerization in addition to the PAS cores. Additionally, the observed PpsR$_{N-Q-PAS1}$ dimer organization appears to resemble a general architecture for the connection of PAS modules with a variety of effector or sensory domains. Furthermore, the long Q-linker helix, which is the result of a fusion of the N-cap of PAS1 with the C-terminal extension of the N-domain, features structural flexibility that plays an important role for signal transduction.

2. Materials and methods

2.1. Cloning of full-length and truncated PpsR variants

The gene encoding PpsR was synthesized with codon optimization for expression in *Escherichia coli* (GeneArt, Invitrogen) and was cloned into the pETM-11 expression vector using *Nco*I and *Not*I restriction sites. The resulting PpsR construct encodes the full-length protein including an N-terminal hexahistidine tag followed by a *Tobacco etch virus* (TEV) protease cleavage site (pET-His-TEV-PpsR$_{full}$). PpsR$_{Q-PAS1}$, PpsR$_{N-Q-PAS1}$ and PpsR$_{\Delta HTH}$ constructs were amplified by PCR using pET-His-TEV-PpsR$_{full}$ as a template and the following primer pairs: 123_fw (5′-ATAT**CCATGG**-GAATTGCAGAAGTTCAGCAGCAGCTG-3′) and 257_rv (5′-ATAT**GCGGCCGC**TTAATCTGCCGGATCAATCTGA-CACAGC-3′) for PpsR$_{Q-PAS1}$, 1_fw (5′-ATAT**CCATGG**-GAATGGGTCTGGCAGGCGGTTC-3′) and 257_rv for PpsR$_{N-Q-PAS1}$, and 1_fw and 379_rv (5′-ATAT**GCGGCCGC**-TTAACGGCTGGTATCACGAACAACC-3′) for PpsR$_{\Delta HTH}$. The start and stop codons introduced by the forward and reverse primers, respectively, are underlined and the *Nco*I and
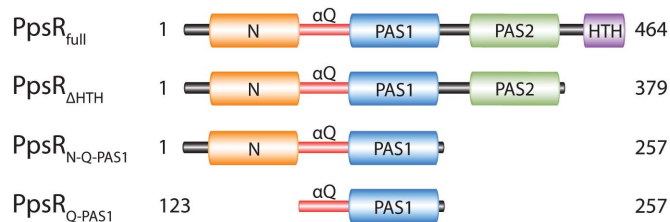

Figure 1
Schematic representation of PpsR$_{full}$ and the crystallized variants. Domains are coloured as follows: N-domain, orange; Q-linker (αQ), red; PAS1, blue; PAS2, green; HTH, purple.

*Not*I restriction sites are highlighted in bold. The truncated PpsR constructs were cloned into the pETM-11 vector as described for pET-His-TEV-PpsR$_{full}$, resulting in pET-His-TEV-PpsR$_{Q-PAS1}$, pET-His-TEV-PpsR$_{N-Q-PAS1}$ and pET-His-TEV-PpsR$_{\Delta HTH}$, respectively.

2.2. Protein expression and purification

Chemically competent *E. coli* BL21 (DE3) cells (Invitrogen) were transformed with the different pET-His-TEV-PpsR plasmid DNAs. Single colonies grown on Luria–Bertani (LB) agar plates containing 50 µg ml^{-1} kanamycin were picked for inoculation of an overnight culture consisting of LB medium supplemented with 30 µg ml^{-1} kanamycin. The overnight culture was used to inoculate 12 l LB medium containing the same antibiotic. The cell cultures were grown to an optical density at 600 nm of 0.4 absorption units (AU) at 37°C, cooled to 18°C and the cells were further grown until the optical density reached 0.8 AU. Protein expression was induced by adding isopropyl *β*-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 m*M*. The cells were harvested by centrifugation 16 h post-induction.
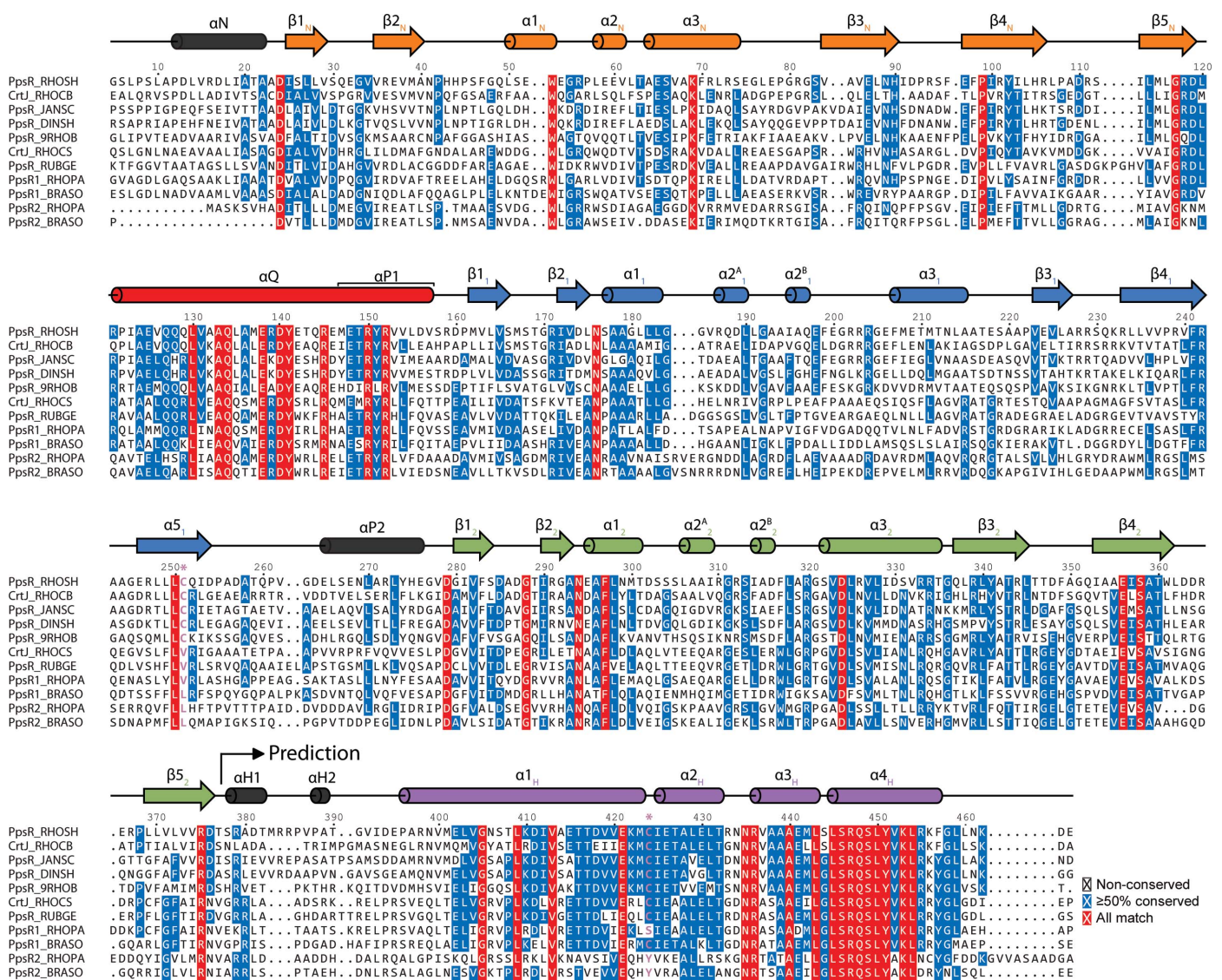

Figure 2
Multiple sequence alignment of PpsR proteins from different organisms. The proteins are listed according to their pairwise sequence identity to PpsR from *R. sphaeroides* in descending order. The aligned proteins are PpsR from *R. sphaeroides* (RHOSH, UniProt accession No. Q9S301), PpsR (CrtJ) from *Rhodobacter capsulatus* (RHOCB, D5ANS9), PpsR from *Jannaschia* sp. (JANSH, Q28W31), PpsR from *Dinoroseobacter shibae* (DINSH, A8LQ24), PpsR from *Thalassiobium* sp. R2A62 (9RHOB, C7DFS5), PpsR (CrtJ) from *Rhodospirillum centenum* (RHOCS, B6ITX1), PpsR from *Rubrivivax gelatinosus* (RUBGE, Q8KRL4), PpsR1 from *Rhodopseudomonas palustris* (RHOPA, Q6N9L3) and *Bradyrhizobium* sp. (BRASO, Q6A567), and PpsR2 from *Rhodopseudomonas palustris* (RHOPA, Q6N9K7) and *Bradyrhizobium* sp. (BRASO, Q8VUB5). The alignment was performed using *Jalview* (Waterhouse *et al.*, 2009) and the *MUSCLE* algorithm (default settings). The secondary-structure elements are drawn according to protomer *A* of the PpsR$_{\Delta HTH}$ structure and a *PSIPRED* (Buchan *et al.*, 2010) prediction of the secondary-structure elements in the HTH motif. The two cysteine residues Cys251 and Cys424, as well as their corresponding residues in related species, are highlighted in pink and marked with an asterisk.

All cell pellets were resuspended in buffer A [10 mM N-cyclohexyl-2-aminoethanesulfonic acid (CHES) pH 9.0, 50 mM imidazole, 300 mM NaCl, 5%(w/v) glycerol] including cOmplete Protease-Inhibitor Cocktail (Roche). The cells were lysed using a microfluidiser (Microfluidics) and the lysates were clarified by ultracentrifugation at 185 500g at 4°C for 1 h.

All hexahistidine-tagged PpsR variants were purified following the same protocol with an additional purification step for PpsR$_{full}$. The cleared supernatant was loaded onto an Ni^{2+}–NTA Superflow (Qiagen) affinity column pre-equilibrated with buffer A. The resin was washed with ten column volumes (CVs) of buffer A and the bound proteins were eluted using 5 CVs of buffer B [10 mM CHES pH 9.0, 200 mM imidazole, 300 mM NaCl, 5%(w/v) glycerol]. PpsR-containing fractions were dialysed against 1 l buffer C [10 mM CHES pH 9.0, 100 mM NaCl, 2 mM dithioerythritol (DTE), 2 mM ethylenediaminetetraacetic acid (EDTA), 5%(w/v) glycerol] at 4°C overnight. In parallel, the hexahistidine tag was removed from the protein using TEV protease with a 1:30 molar ratio of TEV:protein, respectively. The cleaved tag and the histidine-tagged TEV protease were removed from the PpsR-containing solutions by chromatography on Ni^{2+}–NTA resin. Subsequently, the flowthrough of PpsR$_{full}$ only was additionally loaded onto a HiTrap Heparin column (GE Healthcare) equilibrated with buffer D [10 mM CHES pH 9.0, 150 mM NaCl, 5%(w/v) glycerol]. After washing the column with buffer D, the bound protein was eluted in a gradient to 75% buffer D supplemented with 850 mM NaCl over 10 CVs. Fractions containing PpsR$_{full}$ and the Ni^{2+}–NTA flowthrough of the truncated variants were concentrated using centrifugal filter units (Amicon) and were subjected to gel filtration on a Superdex 200 column (GE Healthcare) equilibrated with buffer D. The protein-containing fractions were collected, concentrated, aliquoted, frozen in liquid nitrogen and stored at −80°C.

Selenomethionine-substituted PpsR$_{Q-PAS1}$ and PpsR$_{\Delta HTH}$ proteins were expressed in E. coli BL21 (DE3) according to the protocol of Van Duyne et al. (1993), and protein purification was performed as described above.

2.3. Protein crystallization

Since previous attempts to crystallize PpsR$_{full}$ were unsuccessful, we generated C-terminally truncated PpsR variants and used them in our crystallization experiments. Initial crystallization conditions were identified in sparse-matrix screens. Subsequently, promising crystallization conditions were optimized in a hanging-drop vapour-diffusion setup using 24-well Linbro plates at 293 K. PpsR$_{Q-PAS1}$ crystals with a hexagonal morphology appeared in 2 µl drops consisting of a 1:1 mixture of protein solution (8 mg ml^{-1}) and reservoir solution consisting of 100 mM sodium cacodylate pH 6.5, 1.9%(w/v) polyethylene glycol (PEG) 3000, 200 mM MgCl$_2$ after 1 d. However, individual crystals required up to 10 d to reach their maximal dimensions of ∼60 × 55 × 40 µm. Selenomethionine-substituted PpsR$_{Q-PAS1}$ crystals appeared in a 2 µl 1:1 mixture of protein solution (3.5 mg ml^{-1}) and

reservoir solution consisting of 100 mM 2-(N-morpholino)ethanesulfonic acid (MES) pH 6.5, 3%(w/v) PEG 4000, 200 mM MgCl$_2$, 5 mM DTE and reached their final size after 2 d. For cryoprotection and also for substitution of the cacodylate buffer, crystals obtained from both experiments were incubated in a solution consisting of 100 mM MES pH 6.5, 10%(w/v) PEG 3000, 200 mM MgCl$_2$, 15%(v/v) ethylene glycol for 1 min and were subsequently flash-cooled in liquid nitrogen.

PpsR$_{N-Q-PAS1}$ crystals appeared from a 2 µl 1:1 mixture of protein solution (10 mg ml^{-1}) and reservoir solution consisting of 1.6 M ammonium sulfate, 100 mM MES pH 6.5, 10 mM L-cysteine, 4%(v/v) 1,4-dioxane after 1 d. Individual plate-shaped crystals reached final dimensions of ∼150 × 30 × 5 µm after 5 d but grew in clusters. However, single plates could be isolated from these clusters and were briefly incubated in reservoir solution supplemented with 20%(v/v) ethylene glycol prior to flash-cooling in liquid nitrogen.

Native as well as selenomethionine-substituted PpsR$_{\Delta HTH}$ crystals with a hexagonal morphology started to grow from a mixture of 1.1 µl protein solution (30 mg ml^{-1}) and 0.9 µl reservoir solution consisting of 1 M ammonium sulfate, 100 mM Tris–HCl pH 8.5, 6%(w/v) xylitol, 12%(w/v) glycerol, 1%(v/v) dioxane after 2–3 d. Individual crystals reached their final size of ∼400 × 100 × 80 µm within two weeks. Prior to flash-cooling in liquid nitrogen, native as well as seleno-methionine-substituted PpsR$_{\Delta HTH}$ crystals were incubated for 5 min in buffer consisting of 50 mM sodium phosphate pH 6.8, 12%(w/v) glycerol, 3.0 M sodium malonate pH 7.

2.4. Data collection, structure determination and refinement

Diffraction data were collected from single crystals on beamline X10SA at the Swiss Light Source (Villigen, Switzerland) at 100 K. Indexing, integration and scaling of the data sets was performed using the XDS program suite (Kabsch, 2010). Experimental phases for the PpsR$_{Q-PAS1}$ data were obtained from single-wavelength anomalous diffraction (SAD) experiments on crystals of selenomethionine-substituted protein using AutoSol in PHENIX (Adams et al., 2010). The phases were sufficient to automatically build an initial model using AutoBuild in PHENIX (Terwilliger et al., 2008) using diffraction data to 2.5 Å resolution. This model was further optimized and extended in iterative cycles of manual building using Coot (Emsley et al., 2010) and gradient-driven maximum-likelihood refinement using phenix.refine (Afonine et al., 2012). In the last rounds of model improvement, TLS refinement (Winn et al., 2001) for two individual groups corresponding to residues 143–157 and 158–257 was included. The final model was used for phase extension to the diffraction data at 1.65 Å resolution from isomorphous native PpsR$_{Q-PAS1}$ crystals by rigid-body refinement. This high-resolution model was again optimized by cycles of manual building and refinement similar to those described above.

PpsR$_{N-Q-PAS1}$ was phased by molecular replacement (MR) using Phaser (McCoy et al., 2007) in PHENIX with the PpsR$_{Q-PAS1}$ dimer as the search model. This initial dimer

model consisted of the protomer in the asymmetric unit of the high-resolution PpsR$_{Q-PAS1}$ structure and a symmetry-related molecule (symmetry operator $x - y$, $-y$, $-z$). The phases from MR were of sufficient quality to calculate an electron-density map based on diffraction data to 2.2 Å resolution that enabled building of the missing parts of the polypeptide chain. The structure was then further refined as described above for PpsR$_{Q-PAS1}$ using three TLS groups for each protomer, corresponding to the N-domain (residues 6–120), the Q-linker (residues 121–157) and the PAS1 domain (residues 158–257).

The structure of PpsR$_{\Delta HTH}$ was determined using a combined MR and Se-SAD approach. Statistics of the data collection, structure refinement and model quality of PpsR$_{\Delta HTH}$ and PpsR$_{\Delta HTH}$ Se-SAD have recently been published (Winkler *et al.*, 2013). Initial phases were obtained by MR with *AutoMR* in *PHENIX* using the PpsR$_{N-Q-PAS1}$ structure. However, a unique MR solution was only obtained when PpsR$_{N-Q-PAS1}$ was split into two ensembles, where the first contained the N-terminal region (residues 6–141) in its dimeric form and the second encompassed a monomer of the C-terminal part (residues 145–257). Two molecules of the N-terminal dimer and four molecules of a monomeric C-terminal model were placed in the asymmetric unit, indicating a tetrameric PpsR$_{\Delta HTH}$ assembly. Although the electron-density map after MR was of insufficient quality for model building, the combination of MR phases (from *AutoMR*) with anomalous diffraction data (similar to that described by Schuermann & Tanner, 2003) from a seleno-methionine-substituted PpsR$_{\Delta HTH}$ crystal allowed the identification of the expected 36 maxima in an anomalous difference electron-density map using *AutoSol*. After additional cycles of density modification with *RESOLVE* (Terwilliger, 2000), an interpretable high-quality electron-density map could be calculated using diffraction data to 2.8 Å resolution. A model of PpsR$_{\Delta HTH}$, extending the high-resolution structural models of the N-domain and PAS1 domain, was built and improved in cycles of gradient-driven maximum-likelihood refinement with *phenix.refine* including secondary-structure restraints and manual building using *Coot*. Noncrystallographic symmetry (NCS) restraints were applied to the corresponding PpsR protomers A and D as well as B and C during initial rounds of model building and refinement, but were omitted during the end of the refinement to account for differences between the NCS-related protomers. TLS refinement was included [four TLS groups for each protomer, corresponding to the N-domain (residues 6–120), the Q-linker (residues 121–157) and the PAS1 (158–255) and PAS2 domains (256–378)], during the final refinement cycles.

Statistics of the data collection, structure refinement and model quality of all structures are reported in Table 1. Model quality was analysed using the *MolProbity* (Chen *et al.*, 2010) validation tool as implemented in *PHENIX*. Atomic coordinates of the structures and structure-factor amplitudes have been deposited in the Protein Data Bank as entries 4l9e (PpsR$_{Q-PAS1}$), 4l9f (PpsR$_{Q-PAS1}$ SeMet) and 4l9g (PpsR$_{N-Q-PAS1}$). The PpsR$_{\Delta HTH}$ structure was previously deposited as PDB entry 4hh2 (Winkler *et al.*, 2013). Analysis of the PpsR dimer

Table 1
Data-collection and refinement statistics.

The statistics for PpsR$_{\Delta HTH}$ and PpsR$_{\Delta HTH}$ Se-SAD have recently been published (Winkler *et al.*, 2013). Values in parentheses are for the highest resolution shell.

	PpsR$_{Q-PAS1}$ Se-SAD†	PpsR$_{Q-PAS1}$	PpsR$_{N-Q-PAS1}$
Data collection			
Space group	$P6_1 22$	$P6_1 22$	$P2_1 2_1 2$
Unit-cell parameters (Å)			
a (Å)	50.6	50.4	50.4
b (Å)	50.6	50.4	107.9
c (Å)	163.5	163.3	92.1
$\alpha = \beta$ (°)	90.0	90.0	90.0
γ (°)	120.0	120.0	90.0
Wavelength (Å)	0.9794	0.9765	0.9785
Resolution range (Å)	50–2.5	50–1.65	50–2.2
	(2.60–2.50)	(1.70–1.65)	(2.30–2.20)
Unique reflections	8108 (887)‡	15633 (1299)	26156 (3183)
Completeness (%)	100 (99.8)	99.9 (100)	99.8 (99.7)
Multiplicity	40.8 (38.6)	12.6 (12.7)	4.3 (4.2)
R_{merge} (%)	14.8 (77.0)	4.9 (55.8)	10.4 (63.1)
$\langle I/\sigma(I) \rangle$	26.87 (6.43)	21.06 (4.42)	10.37 (2.38)
Wilson B (Å^2)	36.0	33.2	30.1
Refinement			
Resolution range (Å)	42.3–2.5	43.6–1.65	46.6–2.2
No. of reflections	4783	15631	26156
R_{work}/R_{free}§ (%)	0.1930/0.2373	0.1967/0.2190	0.1767/0.2332
No. of atoms			
Protein	861	894	3862
Ligand/ion	—	—	32
Water	22	85	147
Total	883	979	4041
R.m.s. deviations			
Bond lengths (Å)	0.009	0.006	0.010
Bond angles (°)	1.192	0.966	1.209
B factors			
Protein	41.4	41.7	44.0
Ligand/ion	—	—	65.5
Water	40.5	49.7	43.6
Average	41.3	42.4	44.2
Ramachandran statistics (%)			
Favoured	98.1	99.1	98.2
Allowed	1.9	0.9	1.8
Outliers	0	0	0

† Statistics reported to the cutoff used for anomalous data processing. ‡ Friedel pairs were treated as separate reflections in this case. § R_{free} values were calculated using a randomly assigned 5% of the reflections, which were omitted during refinement.

and tetramer interfaces and the buried surface areas (BSAs) were performed using the *Protein Interfaces, Surfaces and Assemblies* service (*PISA*; v.1.42) at the European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html; Krissinel & Henrick, 2007). Illustrations of the obtained structures were prepared with *PyMOL* v.1.3r1 (Schrödinger).

2.5. Multi-angle light-scattering experiments

Individual PpsR variants were analysed using an HPLC (Waters) setup including a Superdex 200 10/300 GL column (GE Healthcare) connected to a MALS detector (DAWN HELEOS, Wyatt Technology) combined with a refractive-index detector (Waters). The system was equilibrated with buffer D, and 400 µl of a 20 µM protein solution containing the respective PpsR variants was loaded and fractionated at a flow rate of 0.5 ml min^{-1}. Data analysis was performed using
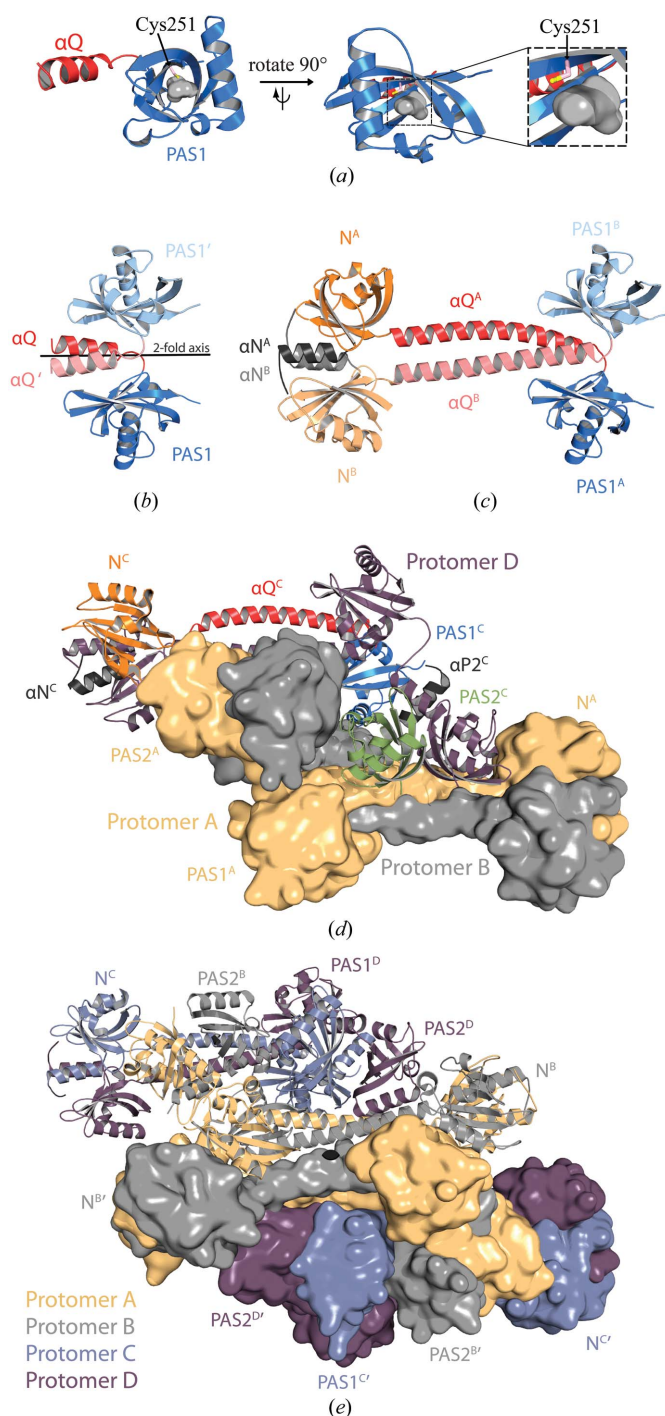
the *ASTRA* software provided by the manufacturer (Wyatt Technology).

2.6. Microscale thermophoresis

The oligomerization of $PpsR_{full}$ and the truncated variants was quantified by microscale thermophoresis (MST) using a Monolith NT.115 (Nanotemper). The proteins were randomly labelled at their amine positions using NT-647 red fluorescent dye (Nanotemper), with approximately one label per monomer, as described in the Monolith NT protein labelling kit RED-NHS provided by the manufacturer (Nanotemper). 1:2 dilution series of unlabelled proteins were prepared over an appropriate concentration range using buffer *D* and were mixed with equivalent volumes of labelled proteins, resulting in final concentrations of 12.5 nM for $PpsR_{full}$ and $PpsR_{\Delta HTH}$ as well as 10 nM for $PpsR_{N-Q-PAS1}$ with respect to the labelled species. Standard treated capillaries (Nanotemper) were used for all experiments. Data for three individual experiments were averaged and evaluated using the quadratic equation of the law of mass action with the constraint of a fixed labelled species concentration.

3. Results

3.1. Structure determination of PpsR$_{\Delta HTH}$

All initial attempts to solve the phase problem of $PpsR_{\Delta HTH}$ (see Fig. 1 for domain architecture) by MR using models of homologous PAS domains failed, likely owing to the low sequence conservation (identities of below 25% with respect to the PAS domains of PpsR). Furthermore, *ab initio* phasing of diffraction data collected from selenomethionine-substituted $PpsR_{\Delta HTH}$ crystals using *HySS* in *PHENIX* (*AutoSol*; Grosse-Kunstleve & Adams, 2003), *autoSHARP* (Vonrhein *et al.*, 2007) and *SHELX* (Sheldrick, 2008) was unsuccessful owing to the weak anomalous signal below 6 Å (Winkler *et al.*, 2013). We therefore applied a multistep divide-and-conquer strategy to successfully determine the structure of the triple PAS protein $PpsR_{\Delta HTH}$. We crystallized PpsR variants that encompass one or two PAS domains, determined their structures and used these models for phasing the diffraction data of longer variants by MR. In a first step, we determined the

Figure 3

(*a*) Cartoon representation of PpsR$_{Q-PAS1}$. The PAS1 domain is coloured blue and the structured C-terminal part of the Q-linker helix is coloured red. Cys251 (pink) located in strand $\beta 5_1$ (Fig. 2) points towards a small cavity (shaped as a grey surface) that is frequently involved in cofactor binding of PAS domains. (*b*) Structure of a parallel PpsR$_{Q-PAS1}$ dimer. The PpsR$_{Q-PAS1}$ protomer is coloured according to (*a*), whereas the symmetry-related molecule (distinguished by a prime) is coloured in equivalent but lighter colours. The crystallographic twofold axis generating the dimer is shown in black. (*c*) Structure of the PpsR$_{N-Q-PAS1}$ dimer. The N-domain of PpsR$_{N-Q-PAS1}$ is coloured orange and its capping helix grey. The remaining structural elements are coloured according to (*a*). Protomer *A* is coloured in saturated colours and protomer *B* in light colours. Individual domains of different protomers are distinguished by superscript letters. (*d*) Structure of the PpsR$_{\Delta HTH}$ tetramer (PDB entry 4hh2; Winkler *et al.*, 2013) formed by two antiparallel dimers. Protomer *C* is coloured according to (*c*) and the additional PAS2 domain is shown in green. The dimer formed by protomers *C* and *D* (coloured purple) is shown as a cartoon representation, whereas that formed by protomers *A* (coloured light orange) and *B* (coloured grey) is depicted as a surface representation. (*e*) The octameric assembly of PpsR$_{\Delta HTH}$ represented by two tetramers generated by a crystallographic twofold axis (Winkler *et al.*, 2013). View along the twofold symmetry axis perpendicular to the plane (black symbol). One tetramer is shown as a cartoon representation and the second as a surface representation. Protomers *A* are coloured in light orange, *B* in grey, *C* in slate and *D* in dark purple, respectively. Individual domains of different protomers are distinguished by superscript letters and labels of the symmetry mate are marked with an additional prime. Note that the N-Q-PAS1 domains of protomers *A* and *B* exclusively stabilize the quaternary assembly.

structure of PpsR$_{Q-PAS1}$ comprising the glutamine-rich linker (Q-linker) and the PAS1 domain (amino acids 123–257; Fig. 1). Phasing was performed by Se-SAD and an initial model was built (Table 1). We finally extended the phases to 1.65 Å resolution using a native data set and refined a high-resolution structure for this variant (Table 1). This model comprised residues 143–257 of one polypeptide chain (Fig. 3a) in the asymmetric unit. The 20 N-terminal residues are disordered and could not be modelled into the electron-density map. The remaining ordered part of the Q-linker region (residues 143–157) forms a helical element that corresponds to a classical PAS N-cap. Together with a molecule related by crystallographic dyad symmetry, PpsR$_{Q-PAS1}$ forms a PAS dimer which features an assembly common to N-cap-comprising PAS domains (Ayers & Moffat, 2008; Fig. 3b).

In a second step, we determined the structure of a PpsR variant comprising the N-terminal domain, the Q-linker and the PAS1 domain (PpsR$_{N-Q-PAS1}$, amino acids 1–257; Fig. 1) at 2.2 Å resolution (Table 1). Phases for PpsR$_{N-Q-PAS1}$ were obtained by MR using the PpsR$_{Q-PAS1}$ dimer as a search model. Since the dimer arrangement of the PAS1 domain was virtually identical in both crystal forms, we obtained excellent phases and could model the remaining parts of the Q-linker and the N-domain into the electron-density map. The final model of PpsR$_{N-Q-PAS1}$ contains a dimer in the asymmetric unit comprising residues 6–246 for protomer A and 8–250 for protomer B (Fig. 3c). Poor electron density was only observed in the $\beta 3_N$–$\beta 4_N$ loop (secondary-structure elements of individual PAS domains are numbered $\beta 1$–$\beta 5$ and $\alpha 1$–$\alpha 3$, with subscripts indicating the corresponding domain: N, N-domain; 1, PAS1 domain; 2, PAS2 domain; Fig. 2) in the N-domain of protomer B.

The structure of the PpsR$_{\Delta HTH}$ construct, which encompasses the N-domain and the PAS1 and PAS2 domains, was finally determined to 2.8 Å resolution using a combination of MR and experimental phases. Individual domains of the PpsR$_{N-Q-PAS1}$ dimer were used as search models for MR and the obtained initial phase information was subsequently used for the identification of Se positions in the PpsR$_{\Delta HTH}$ Se-SAD data (Winkler et al., 2013). Crystallographic analysis of the triple PAS protein PpsR$_{\Delta HTH}$ revealed four polypeptide chains in the asymmetric unit forming an intricate tetrameric assembly (Fig. 3d). Whereas unambiguous and interpretable electron density was observed for most regions of promoters A, B and C, weak electron density was observed for parts of the N-domain and PAS1 domain of protomer D. These parts were modelled taking into account information from the higher resolution structures and protomer A of the PpsR$_{\Delta HTH}$ structure that is related to protomer D by noncrystallographic symmetry. Ambiguous electron density was also observed in the loop connecting the PAS1 and PAS2 domains around residues 258–262 for protomers A, B and C, suggesting high flexibility of this region. However, this region could be modelled in protomer D, and superposition with the other protomers enabled the correct assignment of all PAS2 domains to the individual polypeptide chains. Finally, the five N-terminal residues as well as the C-terminal arginine of all

protomers were disordered and could not be modelled into the electron-density map.

3.2. Structural comparison of PpsR$_{Q-PAS1}$, PpsR$_{N-Q-PAS1}$ and PpsR$_{\Delta HTH}$

The PpsR$_{Q-PAS1}$ structure revealed a typical PAS fold consisting of a five-stranded antiparallel β-sheet flanked by several helices. In addition to the PAS1 core, part of the Q-linker adopts an α-helical conformation corresponding to a PAS N-cap. This is important for dimer formation as observed for other N-cap-containing PAS domains (Ayers & Moffat, 2008). In addition to the evolutionary conservation of such a PAS dimer, analysis of the contact area using the PISA web server (Krissinel & Henrick, 2007) revealed a buried surface area (BSA) of 2590 Å^2 (Fig. 3b) and supported the crystallographic dimer as biologically relevant. The main dimerization interface is formed by the N-cap, its connecting loop to the PAS core (residues 156–161) and $\beta 5_1$ (Supplementary Fig. S1[1]). The N-cap further contributes to PpsR$_{Q-PAS1}$ dimer formation by protecting hydrophobic patches on the PAS core of the symmetry-related molecule (Supplementary Fig. S2), as also reported for a subset of other PAS proteins (Ma et al., 2008). This structure additionally revealed that Cys251, which has been reported to form an intramolecular disulfide bond with Cys424 (Masuda & Bauer, 2002), is located in the middle of $\beta 5_1$, pointing towards a small cavity inside the PAS1 domain (Fig. 3a).

In line with the parallel PpsR$_{Q-PAS1}$ dimer, we observed a dimeric arrangement of the two protomers in the PpsR$_{N-Q-PAS1}$ structure (Fig. 3c). The PpsR$_{N-Q-PAS1}$ dimer has a barbell-like quaternary structure with a protomer crossing along a central imperfect dyad axis. Owing to the protomer crossing, the N-domain and PAS1 domain of a single protomer are located on opposite sides of the dyad axis. The functional relevance of the observed dimer architecture is supported by a large number of contacts, which lead to a total BSA of 6180 Å^2. The PpsR$_{N-Q-PAS1}$ structure also allowed the identification of an additional PAS domain in the N-terminal part of PpsR, which was predicted based on sequence information from some PpsR homologues (Kovács et al., 2005). The PAS core is N-terminally flanked by a helix (αN), which again corresponds to a classical PAS N-cap and constitutes a central part of the N-domain dimer interface (Supplementary Fig. S2), in analogy to the N-cap of PAS1. The highly conserved Q-linkers (Fig. 2) form a ~57 Å long coiled-coil-like structure which connects the N-domain dimers to the PAS1 dimers.

While the arrangement of the structural elements in the PpsR$_{N-Q-PAS1}$ structure is essentially the same as in the PpsR$_{\Delta HTH}$ structure, the latter revealed the importance of the PAS2 domain of PpsR. As observed for the N-domains and the PAS1 domains, the PAS2 domains also form a homodimer and contribute significantly to the overall PpsR dimer interface by the interaction of highly conserved residues located in the PAS2 β-sheets (Glu356 with Arg375) as well as in $\beta 4_2$ and

---

the loop connecting the N-terminal capping helix of PAS2 (αP2) and $\beta1_2$ (Asp279 with Ser358; Supplementary Fig. S1). In addition, the PAS2 N-cap also contributes to dimer formation by interacting with the corresponding N-cap of the second protomer. Notably, His275, which has been proposed to be involved in the coordination of a haem cofactor (Yin *et al.*, 2012), is located on the PAS2 N-cap. However, the PAS2 domains are not only involved in PpsR dimer formation; they also mediate the formation of an intertwined PpsR tetramer formed by two PpsR dimers arranged head to tail (Fig. 3*d*). The PAS2 domains of each dimer interact with the Q-linkers of the other dimer, thereby mediating PpsR tetramer formation. Highly conserved residues of PpsR are involved in this interaction, including elements of helix $\alpha3_2$ as well as $\beta3_2$ of PAS2 and the Q-linker (such as, for example, Arg339 with Glu138 and Leu340 with Gln134; see Fig. 2). Since the PAS2 dimer axis does not align with that of the N-domain and PAS1 dimers, noncrystallographic symmetry is only observed for promoters *A* and *D* as well as for promoters *B* and *C*. Analysis of the contact surfaces using *PISA* supported the relevance of the observed PpsR$_{\Delta HTH}$ tetramer, with a BSA of 26 870 Å^2. However, in addition to the tetramer interface, the Q-linker provides another antiparallel oligomerization interface with a symmetry-related molecule (symmetry operator $-x$, $-x + y$, $-z + 2/3$). This interaction results in the formation of a PpsR octamer (Fig. 3*e*) mediated by contacts of the N-domains, the Q-linkers and the PAS1 domains of the crystallographic and symmetry-related protomers *A* and *B*. The BSA of the octamer interface is 27 160 Å^2, which is comparable to that of the tetramer interface (BSA of 26 870 Å^2). Additional stabilization by the octamer interface is likely to be one reason for the better defined electron density observed for several regions of protomer *A* in comparison to protomer *D*.

Superpositions of identical PAS domains within the protomers of all three PpsR crystal structures showed a very similar three-dimensional conformation of the N-domain and the PAS1 and PAS2 domains. Pronounced differences were only observed for helix $\alpha3_1$ of the PAS1 domain of protomer *A* (PpsR$_{\Delta HTH}$), which is rotated by about 42° compared with all other PAS1 domains (Supplementary Fig. S3). This rotation is

induced by contacts with a symmetry-related molecule in the octamer interface of the PpsR$_{\Delta HTH}$ structure. Comparisons of the N-domain, PAS1 and PAS2 domain dimers, including their N-caps, identified these structures as rigid dimer modules (Fig. 4). Although molecular details of the dimerization interface differ between the various PAS dimers, their dimer architecture is comparable. The N-caps constitute a central part of all dimer interfaces by protecting hydrophobic patches on one side of the PAS β-sheet. However, the interactions between the PAS2 N-caps are not as extensive as those observed for the N-caps of the N-domain and PAS1 domain. In contrast to the N-domain and PAS1 domains, the PAS2 domains make additional dimer contacts *via* their β-sheets (*e.g.* Glu356 with Arg375; Supplementary Fig. S1). Superpositions of all determined PpsR structures based on their PAS1 dimers revealed diverse orientations of the N-terminal helical part of the Q-linker, which results in different relative arrangements of the N-domain and PAS1 domain dimers in the determined PpsR structures (Fig. 5*a*). This is the result of inherent flexibility in a distinct hinge-like Q-linker region (residues 142–146), which is further reflected in subtle differences between the Q-linkers of the symmetrically arranged PpsR$_{\Delta HTH}$ protomers *A* and *D* as well as *B* and *C* (Fig. 5*a*). In contrast, the PAS1 cores and their corresponding N-caps are rigid modules that exhibit an identical conformation in all determined PpsR structures. Since α-helical extensions are not only present at the N-termini but also at the C-termini of some PAS domains (Möglich *et al.*, 2009*b*), this suggests that the long Q-linker helix results from a direct fusion of a C-terminal α-helical extension of the N-domain with the rigid N-cap of the PAS1 domain. Protein-sequence analysis suggests that such long helices, also termed signalling helices (S-helices; Anantharaman *et al.*, 2006), frequently occur in multi-domain signalling proteins and form coiled coils, as observed for the Q-linker of PpsR. These S-helices usually connect an N-terminal sensory domain with a C-terminal effector, or two sensory domains (Anantharaman *et al.*, 2006). Similarly to PpsR, the sensor histidine kinase VicK from *Streptococcus mutans* also possesses a long S-helix, which connects a HAMP and a PAS domain (Wang *et al.*, 2013). The structures of the
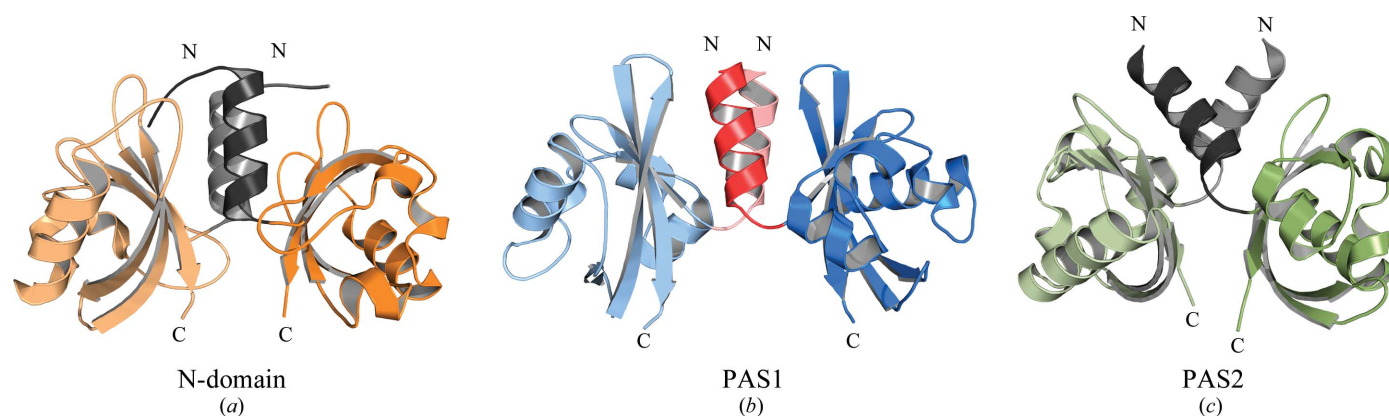

Figure 4
Structures of the three different PpsR PAS domain dimers including their N-caps (from PpsR$_{\Delta HTH}$ protomers *A* and *B*). The PAS core and N-cap of one protomer is coloured according protomer *C* in Fig. 3(*d*), whereas the second protomer is coloured in equivalent but lighter colours. (*a*) N-domain dimer (residues 6–120). (*b*) PAS1 dimer (residues 147–256). (*c*) PAS2 dimer (residues 264–378). The N- and C-termini of each domain are labelled.

PpsR$_{Q\text{-PAS1}}$ dimer (based on the PpsR$_{N\text{-}Q\text{-PAS1}}$ structure) and the corresponding VicK S-helix PAS dimer can be superimposed with an average root-mean-square deviation (r.m.s.d.) of 4.2 Å for 2×131 C$^{\alpha}$ atoms, reflecting the similarities in dimeric arrangements (Fig. 5b). Furthermore, a comparable S-helix PAS organization can also be observed in a truncated structure of the HTR-like protein from *Haloarcula marismortui* (Fig. 5c), which resembles the PpsR$_{Q\text{-PAS1}}$ structure. A search of the Pfam database (Punta *et al.*, 2012) revealed that the S-helix also connects two PAS domains in the HTR-like protein, in analogy to the Q-linker of PpsR. However, since additional PAS and effector domains are present in this protein, the quaternary structure of the full-length protein is difficult to predict.
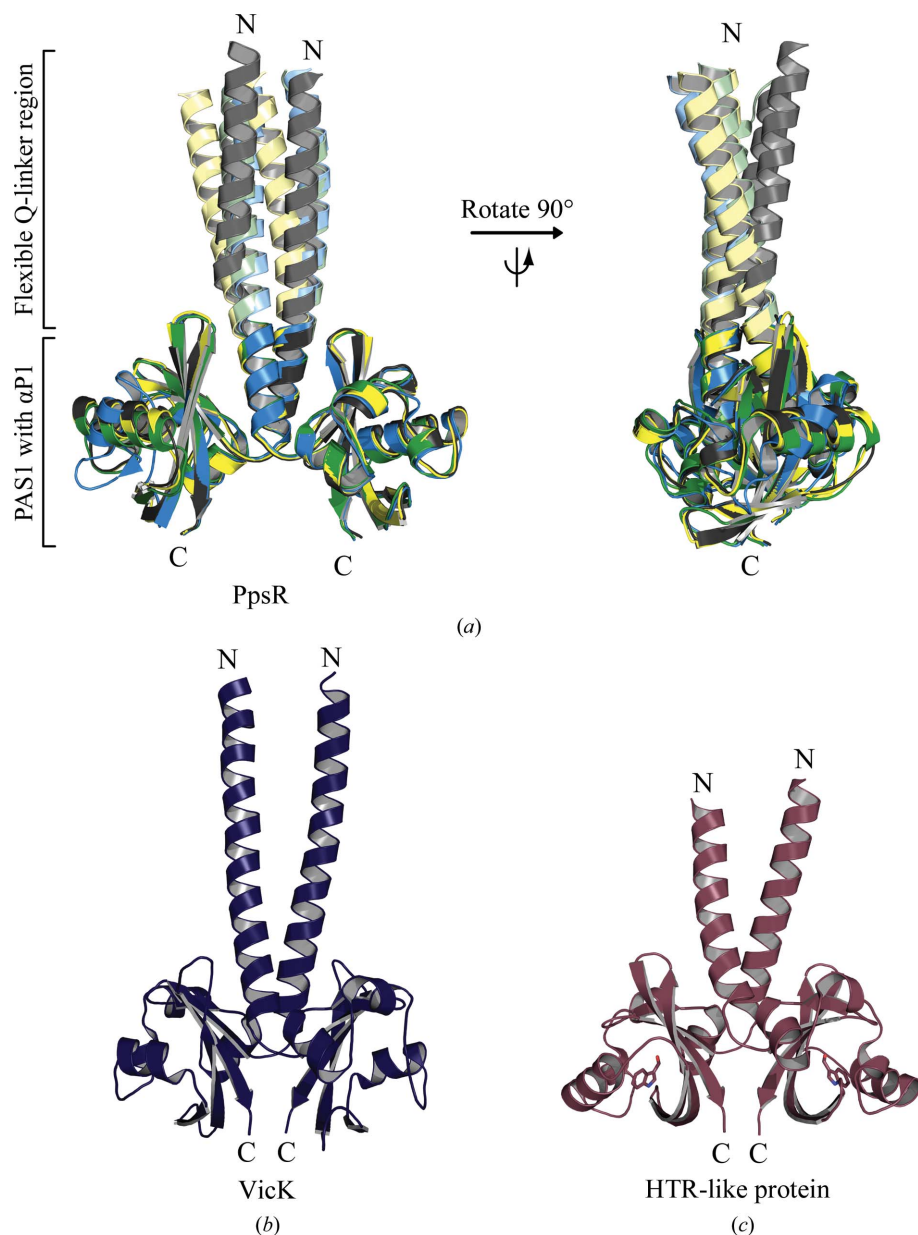
3.3. Oligomerization states of PpsR$_{full}$ and the truncated variants

To better understand the importance of individual PAS domains in oligomerization of PpsR, we determined the molar mass of the crystallized constructs using size-exclusion chromatography coupled to a MALS detector and compared them with PpsR$_{full}$, which has been shown to exist in a dynamic equilibrium of dimers and tetramers in solution (Winkler *et al.*, 2013). Molar masses of 77 and 52 kDa, respectively, were determined for PpsR$_{\Delta HTH}$ and PpsR$_{N\text{-}Q\text{-PAS1}}$ (Fig. 6a). In both cases, this corresponds to approximately 1.8 times the molar mass of a monomer (42 kDa for PpsR$_{\Delta HTH}$ and 29 kDa for PpsR$_{N\text{-}Q\text{-PAS1}}$). The eluting peaks did not show pronounced tailing and no continuous decrease in the molar mass signal was detected as observed for PpsR$_{full}$ (Winkler *et al.*, 2013; Fig. 6a). Together, these findings indicate that both PpsR variants exist predominantly as dimers under the conditions of this experiment. However, the retention time of PpsR$_{\Delta HTH}$ was significantly reduced upon size-exclusion chromatography during protein purification at higher protein concentrations (Fig. 6b) and the elution profile showed peak tailing as observed for PpsR$_{full}$ (Winkler *et al.*, 2013). This finding suggests that at higher concentrations the PpsR$_{\Delta HTH}$ variant forms tetramers similar to PpsR$_{full}$. However, neither tailing nor a reduction in the elution volume of the eluting peak was observed during PpsR$_{N\text{-}Q\text{-PAS1}}$ purification, suggesting that PpsR$_{N\text{-}Q\text{-PAS1}}$ is not able to form tetramers even at high concentrations.

Figure 5
(a) Structural superposition of the PpsR Q-PAS1 dimers of PpsR$_{\Delta HTH}$ (protomers *A* and *B*, blue; protomers *C* and *D*, green), PpsR$_{N\text{-}Q\text{-PAS1}}$ (yellow) as well as the recently published AppA–(PpsR)$_2$ complex structure (black; PDB entry 4hh3; Winkler *et al.*, 2013). The PAS1 dimers and their corresponding N-caps are rigid dimer modules (bright colours) that exhibit an identical conformation in all structures, whereas the corresponding N-terminal helical regions of the Q-linkers (pale colours) show different relative orientations. Even in the AppA–(PpsR)$_2$ complex structure, where the asymmetric binding of AppA to PpsR induces significant structural changes in the conformation of the Q-linker (Winkler *et al.*, 2013), the conformations of the PAS1 and N-domain dimers, including their N-caps, are unchanged. (b) The structure of the S-helix PAS domain dimer of VicK (PDB entry 4i5s; Wang *et al.*, 2013) shows a very similar conformation to that observed for the Q-linkers and PAS1 domains of PpsR. For reasons of clarity, the N-terminal HAMP and the C-terminal S-helix and DHp domains are not shown. (c) The structure of the S-helix PAS domain dimer of HTR-like protein (PDB entry 3bwl; protomers *A* and *B*) also exhibits a comparable conformation to that observed for PpsR and VicK. The PAS domains of the HTR-like protein have an 1*H*-indole-3-carbaldehyde ligand that is shown as stick model.

In contrast to the other studied PpsR variants, $PpsR_{Q-PAS1}$ exists predominantly as a monomer at the concentrations used in this experiment. The average molar mass of 14.3 kDa determined by MALS is in good agreement with the calculated monomer mass of 15.3 kDa.

To analyse and quantify the dimer–tetramer equilibrium of $PpsR_{\Delta HTH}$, we performed microscale thermophoresis measurements (Fig. 7). A K_d value of 0.9 $\pm$ 0.2 µM was determined for 2 $(PpsR_{\Delta HTH})_2 \rightleftharpoons (PpsR_{\Delta HTH})_4$ based on the total dimer concentration. This value is in good agreement with the previously reported K_d value of 0.9 $\pm$ 0.3 µM determined for $PpsR_{full}$ (Winkler et al., 2013). Although the K_d

values for the dimer–tetramer equilibrium of $PpsR_{full}$ and $PpsR_{\Delta HTH}$ are comparable, MALS measurements at comparable concentrations showed such an equilibrium only for $PpsR_{full}$ and not for $PpsR_{\Delta HTH}$. This suggests that the HTH domain of PpsR influences the kinetic stability of the tetramer. In addition to the measurements for $PpsR_{\Delta HTH}$, we also performed MST measurements for the $PpsR_{N-Q-PAS1}$ and $PpsR_{Q-PAS1}$ variants. In agreement with the observation of exclusively dimeric $PpsR_{N-Q-PAS1}$ in our MALS experiments, we quantified the 2 $PpsR_{N-Q-PAS1} \rightleftharpoons (PpsR_{N-Q-PAS1})_2$ transition with a K_d value of 0.25 $\pm$ 0.05 µM. In contrast, for $PpsR_{Q-PAS1}$ we could not detect dimerization over the concentration range used in the MST experiments. Nevertheless, the concentrations used, for example, for crystallization also supported the formation of a biologically relevant $PpsR_{Q-PAS1}$ dimer.

4. Discussion

We provide a detailed structural characterization of PpsR from R. sphaeroides, which reveals the importance of individual PAS domains and their α-helical extensions for PpsR oligomerization. Whereas the implications of the $PpsR_{\Delta HTH}$ structure in the context of the $AppA–PpsR_2$ interaction and the regulation of gene expression in R. sphaeroides have recently been described (Winkler et al., 2013); here, we discuss the structural aspects of PpsR, which enable a better understanding of the architecture and functionality of multi-PAS proteins in general. The $PpsR_{\Delta HTH}$ structure is the first protein structure with three classical PAS domains present on the same polypeptide chain and reveals an intertwined tetrameric arrangement formed by two head-to-tail dimers. Each dimer is formed by two head-to-head aligned PpsR protomers featuring extensive contacts between all domains. The highly
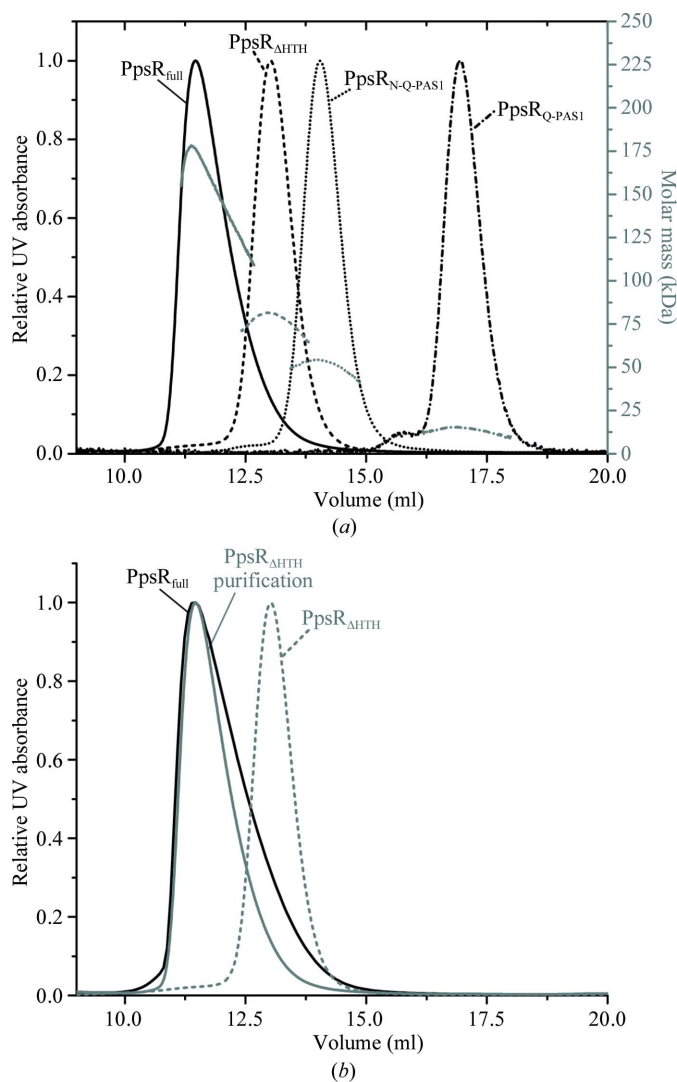

Figure 6
(a) Multi-angle light-scattering detection of $PpsR_{full}$ (straight line), $PpsR_{\Delta HTH}$ (dashed line), $PpsR_{N-Q-PAS1}$ (dotted line) and $PpsR_{Q-PAS1}$ (dashed/dotted line) fractionated by size-exclusion chromatography (individual traces are normalized according to the absorbance at 280 nm). The MALS-derived calculated molar-mass signals (grey) indicated on the right y axis yield average values of 150, 77, 52 and 14.3 kDa, respectively. (b) Absorbance traces of $PpsR_{full}$ (black line) and $PpsR_{\Delta HTH}$ (grey dashed line) at 280 nm from the multi-angle light-scattering experiment in comparison to the absorbance trace of $PpsR_{\Delta HTH}$ (grey solid line) during preparative purification. The $PpsR_{\Delta HTH}$ concentration during preparative purification was about 30-fold higher compared with the multi-angle light-scattering experiment.
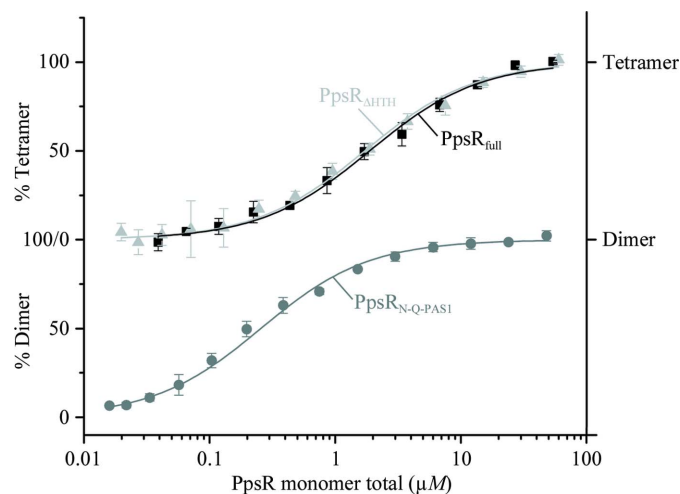

Figure 7
Microscale thermophoresis measurements of the dimer–tetramer equilibria of $PpsR_{full}$ (black line and squares) and $PpsR_{\Delta HTH}$ (light grey line and triangles) result in K_d values of 0.9 $\pm$ 0.3 and 0.9 $\pm$ 0.2 µM, respectively, based on the total dimer concentration. Measurements of the $PpsR_{N-Q-PAS1}$ monomer–dimer equilibrium (dark grey line and filled circles) result in a K_d value of 0.25 $\pm$ 0.05 µM, based on the total monomer concentration. Error bars represent the standard deviation of three individual experiments.

conserved Q-linker connects the N-domain and the PAS1 domain and forms a long coiled-coil-like structure, which is involved in formation of the PpsR dimer as well as the tetramer. As proposed previously, the N-domain also significantly contributes to PpsR oligomer formation and/or integrity (Yamazaki *et al.*, 2008), most likely by stabilizing the extended Q-linker helix as shown by the disordered N-terminal Q-linker part of the PpsR$_{Q-PAS1}$ structure. PpsR constructs lacking the N-domain are unable to bind to DNA *in vivo* (Gomelsky *et al.*, 2000) and *in vitro* (Yamazaki *et al.*, 2008), which highlights the central importance of PpsR oligomerization. In addition to its role in PpsR dimer and tetramer formation, the Q-linker provides an additional oligomerization interface that allows PpsR octamer formation (Fig. 3*e*). Such an octameric assembly has been reported to play a role in the binding of PpsR to its palindromic DNA target sequence (Jaubert *et al.*, 2004; Winkler *et al.*, 2013). Taken together, this highlights the involvement of the α-helical PAS extensions in the oligomerization and functionality of multi-PAS proteins.

This is also supported by MALS and MST experiments of PpsR$_{full}$ and PpsR$_{\Delta HTH}$, which clearly show PpsR tetramer formation in solution. However, the crystal structure does not allow differentiation between tetramers formed by protomers *AB* and *CD* or by two *AB* pairs (A_2B_2) which are related by a crystallographic twofold axis, since the individual BSAs of the two tetramer assemblies are in comparable ranges. Nevertheless, the *ABCD* tetramer is most likely, since the truncated PpsR$_{N-Q-PAS1}$ variant, which provides the entire interface for an A_2B_2 tetramer, is not able to form tetramers. In contrast, PpsR$_{full}$ and PpsR$_{\Delta HTH}$ show the characteristic dimer–tetramer equilibrium, supporting a role of PAS2 in tetramer formation. Further support for the functional relevance of the *ABCD* tetramer is provided by interactions between the highly conserved helix α3$_2$ and the Q-linker (Fig. 2). It has been demonstrated that the Q-linker is not only involved in PpsR oligomerization but also serves as a binding site for the regulatory protein AppA (Winkler *et al.*, 2013) that enables the blue-light- and oxygen-dependent transcriptional control of PpsR-regulated genes (Gomelsky & Kaplan, 1998; Masuda & Bauer, 2002; Braatsch *et al.*, 2002; Jäger *et al.*, 2007; Han *et al.*, 2007). AppA interacts *via* its four-helix bundle with the PAS2 binding region of the Q-linker to form an AppA–PpsR$_2$ complex, thereby preventing PpsR tetramer formation. Hydrogen–deuterium exchange experiments comparing this dimeric state of AppA-complexed PpsR and PpsR predominantly in its tetrameric form further support the biological relevance of the crystallographic tetramer observed in the PpsR$_{\Delta HTH}$ structure (Winkler *et al.*, 2013).

However, the DNA binding of PpsR and its homologues is not exclusively regulated by protein–protein interactions. A variety of regulatory mechanisms involving the partially conserved cysteine (Cys424; Fig. 2) located close to the HTH motif have been proposed (Masuda *et al.*, 2002; Masuda & Bauer, 2002; Jaubert *et al.*, 2004; Yin *et al.*, 2012; Cheng *et al.*, 2012). Depending on the species, this cysteine has been reported to be involved in intermolecular (Jaubert *et al.*, 2004) or intramolecular (Masuda & Bauer, 2002; Masuda *et al.*, 2002)

disulfide-bond formation, cofactor binding (Yin *et al.*, 2012) or as the target for oxidative modifications (Cheng *et al.*, 2012). Early studies on PpsR from *R. sphaeroides* and the PpsR homologue CrtJ from *R. capsulatus* suggested that redox-dependent DNA binding can be regulated *via* the formation of an intramolecular disulfide bond between the two cysteines located in the HTH and the PAS1 domain (Masuda & Bauer, 2002; Masuda *et al.*, 2002; Cho *et al.*, 2004). However, considering the location of Cys251 within the PAS1 domain, the formation of an intramolecular disulfide bond involving Cys251 appears to be unlikely. Disulfide-bond formation would require significant structural changes of PAS1 and would most probably result in disruption of the PAS fold. Interestingly, Cys251 points into a small cavity inside the PAS1 domain where cofactors can be covalently or noncovalently bound in a variety of PAS proteins.

Like PpsR, many PAS proteins contain multiple PAS domains that either bind cofactors and sense environmental stimuli or exclusively serve as protein-interaction or signal transmission modules. Therefore, multi-PAS domain-mediated protein–protein interaction and oligomerization, as observed for PpsR, appears to be a common feature of many PAS-containing proteins. The PpsR$_{\Delta HTH}$ structure highlights this function and shows that not only the PAS core regions, but also the N- and C-terminal α-helical extensions, can significantly contribute to protein oligomer formation. Structural studies of PAS proteins such as *Rm*FixL (Miyatake *et al.*, 2000), *Ec*Dos (Kurokawa *et al.*, 2004), *Av*NifL (PAS1; Key *et al.*, 2007), *Ec*TyrR (Verger *et al.*, 2007), *Np*STHK (Ma *et al.*, 2008), *Mt*Rv1364c (Jaiswal *et al.*, 2010) and *Sm*VicK (Wang *et al.*, 2013) revealed that structures containing N-caps frequently form homodimers that exhibit a parallel arrangement similar to the three PAS domains of PpsR. In contrast, structures lacking N-terminal helices [*Bj*FixLH (Ayers & Moffat, 2008), *Bs*YtvA-LOV (Möglich & Moffat, 2007) and *Cr*Phot-LOV1 (Fedorov *et al.*, 2003)] exhibit diverse quaternary assemblies (Ayers & Moffat, 2008). The important role of N-caps in PAS dimer formation is also supported by the crystal structure of the heterodimeric mouse CLOCK–BMAL1 transcriptional activator complex (Huang *et al.*, 2012). The PAS-A domains of CLOCK and BMAL1, which both possess N-caps, heterodimerize in an arrangement similar to the PAS domains of PpsR. In contrast, their PAS-B domains as well as the PAS-B domains of HIF-2α and ARNT (Scheuermann *et al.*, 2009), which lack N-caps, dimerize *via* different mechanisms. Additionally, studies of the PAS domain of *Np*STHK (termed H-NOXA) show that the deletion of seven N-cap residues disrupts the parallel dimer arrangement and results in a flipped, most likely nonphysiological, dimer (Ma *et al.*, 2008). Similarly, structural and functional studies on KinA have also shown the importance of the N-terminal PAS$_B$ extension for histidine kinase activity (Winnen *et al.*, 2013). Even mutations within the N-cap can influence the quaternary structure of PAS proteins and thus have a strong impact on protein function. It has recently been shown for the CLOCK–BMAL1 PAS-A system that mutation of two residues at the interface of the N-cap dimer interferes with heterodimeriza-

tion and significantly reduces transactivation activity (Huang *et al.*, 2012). Also, single substitutions within the N-cap of the histidine kinases *Sp*MicB (Echenique & Trombe, 2001), *Ec*DcuS (PAS$_C$; Monzel *et al.*, 2013) and the histidine kinase-like protein *Av*NifL (PAS1; Little *et al.*, 2012) can affect protein function, most likely owing to impaired dimerization of the proteins. Additionally, N-cap mutations can affect cofactor binding and thus PAS-mediated signalling, as shown for the aerotaxis receptor *Ec*Aer (Watts *et al.*, 2006). The important function of N-caps in signal transduction and protein oligomerization is further supported by studies on the LOV protein VIVID from *Neurospora crassa* (VVD). Blue-light illumination of VVD results in a series of structural changes that eventually affect the N-cap conformation and thus induce VVD dimerization (Zoltowski *et al.*, 2007; Zoltowski & Crane, 2008; Vaidya *et al.*, 2011). However, not all PAS domains that possess N-caps dimerize in an assembly as observed for PpsR and the proteins mentioned above. The sensory histidine kinases DctB of *Sinorhizobium meliloti* (Zhou *et al.*, 2008) and DcuS (PAS$_P$) of *E. coli* (Cheung & Hendrickson, 2008), for example, mainly dimerize *via* a long helix preceding the classical N-cap, whereas the core regions of their PAS domains do not contribute to dimer formation. However, in the cases of DctB and DcuS (PAS$_P$) the PAS N-caps also help to stabilize the homodimer in its ligand-free and ligand-bound state, respectively.

Helices flanking the PAS core are observed at the N-terminus but also at the C-terminus of PAS domains. Protein-sequence analysis suggests that these α-helical extensions are frequently much longer than classical N-caps and often form coiled coils (Anantharaman *et al.*, 2006; Möglich *et al.*, 2009*b*) similar to the Q-linker of PpsR. These long helices, termed S-helices (Anantharaman *et al.*, 2006), occur in a wide range of multi-domain signalling proteins and frequently link PAS to effector domains, PAS and sensory modules as well as tandem PAS domains. It has been proposed that the coiled-coil S-helices play an important role in signal propagation from sensory domains to effector domains or other sensory domains (Anantharaman *et al.*, 2006; Möglich *et al.*, 2009*a*; Hao *et al.*, 2011; Little *et al.*, 2012). Additionally, they can be involved in the control of protein oligomerization by serving as binding sites for regulatory proteins, as shown for the binding of AppA to the Q-linker of PpsR (Winkler *et al.*, 2013). Upon binding, AppA induces an overall asymmetry in the AppA–PpsR$_2$ complex that results in structural changes of the N-terminal part of one Q-linker, whereas the conformation of the C-terminal part, which corresponds to one N-cap of the PAS1 dimer, and the N-domain dimer remain unaffected and resemble that observed in other PpsR structures (Fig. 5*a*). This suggests that the PAS domain dimers together with their N-caps function as rigid modules. The flexibility of a defined hinge region within the Q-linkers, however, supports the finding that the long Q-linker helices observed in the PpsR$_{\Delta HTH}$ structure result from a direct fusion of the C-terminal extensions of the N-domains with the N-cap of the PAS1 domains. Similar to the Q-linker of PpsR, *Sm*VicK also possesses a long S-helix that originates from a fusion of a

HAMP helix (as commonly found in **h**istidine kinases, **a**denylyl cyclases, **m**ethyl-accepting chemotaxis proteins and **p**hosphatases) with the N-cap of a PAS domain (Wang *et al.*, 2013). The S-helix–PAS dimer arrangement observed for *Sm*VicK is nearly identical to that observed for the PpsR$_{\Delta HTH}$ Q-PAS1 dimer (Fig. 5*b*). Additionally, a similar S-helix PAS dimer arrangement is also observed for the HTR-like protein from *H. marismortui* (Fig. 5*c*; PDB entry 3bwl; Midwest Center for Structural Genomics, unpublished work). Although the N-terminal region of the S-helix and the adjacent PAS domain are not part of the HTR-like protein crystal structure, it is likely that the dimer arrangement of tandem PAS domains featuring characteristic N-caps is similar to that observed for PpsR$_{N-Q-PAS1}$. Therefore, the PAS–S-helix–PAS dimer organization observed for PspR might not only be a conserved dimerization motif in a variety of multi-PAS containing signalling proteins, but might also be a general architecture for the connection of PAS domains to a variety of effector and sensory domains such as REC, GAF, HAMP, MEDS or PocR domains. For the engineered light-regulated sensor histidine kinase YF1, where an S-helix connects a LOV light-sensor to a dimerization/histidine phosphotransfer domain, it has been proposed that the sensed signal is transmitted along the coiled-coil S-helices in the form of torque or helical rotation (Möglich *et al.*, 2009*a*; Diensthuber *et al.*, 2013). Additionally, the same mechanism has also been proposed for the sensory protein *Av*NifL (Little *et al.*, 2012), which exhibits a PAS–S-helix–PAS dimer organization as observed for PpsR. Also, upon AppA–PpsR$_2$ complex formation a pronounced rotation of the PAS1 domain dimer along the Q-linker axis was observed (Winkler *et al.*, 2013). Therefore, it can be suggested that in many PAS-containing proteins signals are transmitted along S-helices *via* such a mechanism.

However, C-terminal PAS extensions do not always extend from the PAS core and/or form coiled coils. They can also interact with the surface of the PAS β-sheet, as observed for the Jα helix in the *Avena sativa* phototropin 1 PAS B (LOV2) domain (Halavaty & Moffat, 2007). Also in the case of LOV2, the C-terminal α-helix is proposed to play an important role in signal transduction by transmitting light-induced structural changes from the PAS core to the kinase domain (Harper *et al.*, 2003, 2004; Halavaty & Moffat, 2007). Additionally, C-terminal PAS extensions can interact with the helical part of the PAS core, as observed for the αE helix of the *Drosophila* protein PERIOD (Yildiz *et al.*, 2005; King *et al.*, 2011) and its three mouse homologues mPER1 (Kucera *et al.*, 2012), mPER2 (Hennig *et al.*, 2009) and mPER3 (Kucera *et al.*, 2012). Taken together, although the structural details of the N-terminal and C-terminal PAS extensions can vary between different proteins, they play an important role in PAS-mediated protein oligomerization and signal transduction.

To better understand signal transduction from PAS to effector domains, as well as the interplay between multiple PAS domains and their N-terminal and C-terminal α-helical extensions in protein oligomerization, structures of full-length proteins are required. However, owing to the inherent flexibility of multi-domain PAS signalling proteins, their

crystallization and structure determination is challenging. For the first time, the determined PpsR structures reveal how three PAS domains within a single peptide chain, together with their N-terminal and C-terminal α-helical extensions, enable the formation of multiple oligomeric states (dimer, tetramer, octamer). Additionally, the PAS–S-helix–PAS dimer organization observed for PpsR appears to be a common architecture for the connection of PAS domains to a variety of effector and sensory domains in multi-domain signalling proteins. The signals sensed by such PAS domains are transmitted *via* S-helices to effector or other sensory domains, thus enabling PAS domains to regulate the activity of a variety of different effector domains without directly interacting with them. Considering that the oligomerization states of PAS domain-containing proteins can be regulated by their cellular abundance or regulatory proteins, changes in quaternary structure as well as the modulation of the homo- or hetero-oligomerization affinity by cofactor binding appear to be general aspects of PAS domains and to be an essential feature for efficient signal processing and/or transduction.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.

Ali, M. H. & Imperiali, B. (2005). *Bioorg. Med. Chem.* **13**, 5013–5020.

Anantharaman, V., Balaji, S. & Aravind, L. (2006). *Biol. Direct*, **1**, 25.

Ayers, R. A. & Moffat, K. (2008). *Biochemistry*, **47**, 12078–12086.

Braatsch, S., Gomelsky, M., Kuphal, S. & Klug, G. (2002). *Mol. Microbiol.* **45**, 827–836.

Buchan, D. W., Ward, S. M., Lobley, A. E., Nugent, T. C., Bryson, K. & Jones, D. T. (2010). *Nucleic Acids Res.* **38**, W563–W568.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Cheng, Z., Wu, J., Setterdahl, A., Reddie, K., Carroll, K., Hammad, L. A., Karty, J. A. & Bauer, C. E. (2012). *Mol. Microbiol.* **85**, 734–746.

Cheung, J. & Hendrickson, W. A. (2008). *J. Biol. Chem.* **283**, 30256–30265.

Cho, S.-H., Youn, S.-H., Lee, S.-R., Yim, H.-S. & Kang, S.-O. (2004). *Microbiology*, **150**, 697–705.

Diensthuber, R. P., Bommer, M., Gleichmann, T. & Möglich, A. (2013). *Structure*, **21**, 1127–1136.

Echenique, J. R. & Trombe, M.-C. (2001). *J. Bacteriol.* **183**, 4599–4608.

Elsen, S., Jaubert, M., Pignol, D. & Giraud, E. (2005). *Mol. Microbiol.* **57**, 17–26.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M. & Hegemann, P. (2003). *Biophys. J.* **84**, 2474–2482.

Gomelsky, M., Horne, I. M., Lee, H.-J., Pemberton, J. M., McEwan, A. G. & Kaplan, S. (2000). *J. Bacteriol.* **182**, 2253–2261.

Gomelsky, M. & Kaplan, S. (1998). *J. Biol. Chem.* **273**, 35319–35325.

Grosse-Kunstleve, R. W. & Adams, P. D. (2003). *Acta Cryst.* D**59**, 1966–1973.

Halavaty, A. S. & Moffat, K. (2007). *Biochemistry*, **46**, 14001–14009.

Han, Y., Meyer, M. H. F., Keusgen, M. & Klug, G. (2007). *Mol. Microbiol.* **64**, 1090–1104.

Hao, N., Whitelaw, M. L., Shearwin, K. E., Dodd, I. B. & Chapman-Smith, A. (2011). *Nucleic Acids Res.* **39**, 3695–3709.

Harper, S. M., Christie, J. M. & Gardner, K. H. (2004). *Biochemistry*, **43**, 16184–16192.

Harper, S. M., Neil, L. C. & Gardner, K. H. (2003). *Science*, **301**, 1541–1544.

Hefti, M. H., Françoijs, K. J., de Vries, S. C., Dixon, R. & Vervoort, J. (2004). *Eur. J. Biochem.* **271**, 1198–1208.

Hennig, S., Strauss, H. M., Vanselow, K., Yildiz, Ö., Schulze, S., Arens, J., Kramer, A. & Wolf, E. (2009). *PLoS Biol.* **7**, e94.

Henry, J. T. & Crosson, S. (2011). *Annu. Rev. Microbiol.* **65**, 261–286.

Herrou, J. & Crosson, S. (2011). *Nature Rev. Microbiol.* **9**, 713–723.

Huang, N., Chelliah, Y., Shan, Y., Taylor, C. A., Yoo, S.-H., Partch, C., Green, C. B., Zhang, H. & Takahashi, J. S. (2012). *Science*, **337**, 189–194.

Jäger, A., Braatsch, S., Haberzettl, K., Metz, S., Osterloh, L., Han, Y. & Klug, G. (2007). *J. Bacteriol.* **189**, 2274–2282.

Jaiswal, R. K., Manjeera, G. & Gopal, B. (2010). *Biochem. Biophys. Res. Commun.* **398**, 342–349.

Jaubert, M., Zappa, S., Fardoux, J., Adriano, J. M., Hannibal, L., Elsen, S., Lavergne, J., Verméglio, A., Giraud, E. & Pignol, D. (2004). *J. Biol. Chem.* **279**, 44407–44416.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kewley, R. J., Whitelaw, M. L. & Chapman-Smith, A. (2004). *Int. J. Biochem. Cell Biol.* **36**, 189–204.

Key, J., Hefti, M., Purcell, E. B. & Moffat, K. (2007). *Biochemistry*, **46**, 3614–3623.

King, H. A., Hoelz, A., Crane, B. R. & Young, M. W. (2011). *J. Mol. Biol.* **413**, 561–572.

Kovács, A. T., Rákhely, G. & Kovács, K. L. (2005). *Res. Microbiol.* **156**, 619–625.

Krell, T., Lacal, J., Busch, A., Silva-Jiménez, H., Guazzaroni, M. E. & Ramos, J. L. (2010). *Annu. Rev. Microbiol.* **64**, 539–559.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Kucera, N., Schmalen, I., Hennig, S., Öllinger, R., Strauss, H. M., Grudziecki, A., Wieczorek, C., Kramer, A. & Wolf, E. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 3311–3316.

Kurokawa, H., Lee, D. S., Watanabe, M., Sagami, I., Mikami, B., Raman, C. S. & Shimizu, T. (2004). *J. Biol. Chem.* **279**, 20186–20193.

Little, R., Slavny, P. & Dixon, R. (2012). *PLoS One*, **7**, e46651.

Ma, X., Sayed, N., Baskaran, P., Beuve, A. & van den Akker, F. (2008). *J. Biol. Chem.* **283**, 1167–1178.

Masuda, S. & Bauer, C. E. (2002). *Cell*, **110**, 613–623.

Masuda, S., Dong, C., Swem, D., Setterdahl, A. T., Knaff, D. B. & Bauer, C. E. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 7078–7083.

Matthews, J. M. & Sunde, M. (2012). *Adv. Exp. Med. Biol.* **747**, 1–18.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Miyatake, H., Mukai, M., Park, S.-Y., Adachi, S., Tamura, K., Nakamura, H., Nakamura, K., Tsuchiya, T., Iizuka, T. & Shiro, Y. (2000). *J. Mol. Biol.* **301**, 415–431.

Möglich, A., Ayers, R. A. & Moffat, K. (2009a). *J. Mol. Biol.* **385**, 1433–1444.

Möglich, A., Ayers, R. A. & Moffat, K. (2009b). *Structure*, **17**, 1282–1294.

Möglich, A. & Moffat, K. (2007). *J. Mol. Biol.* **373**, 112–126.

Monzel, C., Degreif-Dünnwald, P., Gröpper, C., Griesinger, C. & Unden, G. (2013). *Microbiologyopen*, **2**, 912–927.

Moskvin, O. V., Gomelsky, L. & Gomelsky, M. (2005). *J. Bacteriol.* **187**, 2148–2156.

Pellequer, J.-L., Wager-Smith, K. A., Kay, S. A. & Getzoff, E. D. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5884–5890.

Penfold, R. J. & Pemberton, J. M. (1991). *Curr. Microbiol.* **23**, 259–263.

Punta, M. et al. (2012). *Nucleic Acids Res.* **40**, D290–D301.

Scheuermann, T. H., Tomchick, D. R., Machius, M., Guo, Y., Bruick, R. K. & Gardner, K. H. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 450–455.

Schuermann, J. P. & Tanner, J. J. (2003). *Acta Cryst.* D**59**, 1731–1736.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Taylor, B. L. & Zhulin, I. B. (1999). *Microbiol. Mol. Biol. Rev.* **63**, 479–506.

Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst.* D**64**, 61–69.

Vaidya, A. T., Chen, C.-H., Dunlap, J. C., Loros, J. J. & Crane, B. R. (2011). *Sci. Signal.* **4**, ra50.

Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.

Verger, D., Carr, P. D., Kwok, T. & Ollis, D. L. (2007). *J. Mol. Biol.* **367**, 102–112.

Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.

Wang, C., Sang, J., Wang, J., Su, M., Downey, J. S., Wu, Q., Wang, S., Cai, Y., Xu, X., Wu, J., Senadheera, D. B., Cvitkovitch, D. G., Chen, L., Goodman, S. D. & Han, A. (2013). *PLoS Biol.* **11**, e1001493.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. (2009). *Bioinformatics*, **25**, 1189–1191.

Watts, K. J., Sommer, K., Fry, S. L., Johnson, M. S. & Taylor, B. L. (2006). *J. Bacteriol.* **188**, 2154–2162.

Winkler, A., Heintz, U., Lindner, R., Reinstein, J., Shoeman, R. L. & Schlichting, I. (2013). *Nature Struct. Mol. Biol.* **20**, 859–867.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.

Winnen, B., Anderson, E., Cole, J. L., King, G. F. & Rowland, S. L. (2013). *J. Bacteriol.* **195**, 2349–2358.

Yamazaki, Y., Fukusumi, H., Kamikubo, H. & Kataoka, M. (2008). *Photochem. Photobiol.* **84**, 839–844.

Yildiz, Ö., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S., Urbanke, C., Sassone-Corsi, P. & Wolf, E. (2005). *Mol. Cell*, **17**, 69–82.

Yin, L., Dragnea, V. & Bauer, C. E. (2012). *J. Biol. Chem.* **287**, 13850–13858.

Zhou, Y.-F., Nan, B., Nan, J., Ma, Q., Panjikar, S., Liang, Y.-H., Wang, Y. & Su, X.-D. (2008). *J. Mol. Biol.* **383**, 49–61.

Zoltowski, B. D. & Crane, B. R. (2008). *Biochemistry*, **47**, 7012–7019.

Zoltowski, B. D., Schwerdtfeger, C., Widom, J., Loros, J. J., Bilwes, A. M., Dunlap, J. C. & Crane, B. R. (2007). *Science*, **316**, 1054–1057.