The FEBS Journal

FEBS PRESS science publishing by scientists

Structural basis for high-affinity adipate binding to AdpC (RPA4515), an orphan periplasmic-binding protein from the tripartite tricarboxylate transporter (TTT) family in *Rhodopseudomonas palustris*

Leonardo T. Rosa, Samuel R. Dix, John B. Rafferty and David J. Kelly

Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, UK

Keywords
adipic acid; C6-dicarboxylate; *Rhodopseudomonas palustris*; substrate-binding protein; tripartite tricarboxylate transporter

Correspondence
D. J. Kelly, Department of Molecular Biology and Biotechnology, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK
Tel: +44 0114 222 4414
E-mail: d.kelly@sheffield.ac.uk

(Received 2 August 2017, revised 27 September 2017, accepted 19 October 2017)

doi:10.1111/febs.14304

The tripartite tricarboxylate transporter (TTT) family is a poorly characterised group of prokaryotic secondary solute transport systems, which employ a periplasmic substrate-binding protein (SBP) for initial ligand recognition. The substrates of only a small number of TTT systems are known and very few SBP structures have been solved, so the mechanisms of SBP–ligand interactions in this family are not well understood. The SBP RPA4515 (AdpC) from *Rhodopseudomonas palustris* was found by differential scanning fluorescence and isothermal titration calorimetry to bind aliphatic dicarboxylates of a chain length of six to nine carbons, with K_D values in the μM range. The highest affinity was found for the C6-dicarboxylate adipate (1,6-hexanedioate). Crystal structures of AdpC, either adipate or 2-oxoadipate bound, revealed a lack of positively charged amino acids in the binding pocket and showed that water molecules are involved in bridging hydrogen bonds to the substrate, a conserved feature in the TTT SBP family that is distinct from other types of SBP. In AdpC, both of the ligand carboxylate groups and a linear chain conformation are needed for coordination in the binding pocket. RT-PCR showed that *adpC* expression is upregulated by low environmental adipate concentrations, suggesting adipate is a physiologically relevant substrate but as *adpC* is not genetically linked to any TTT membrane transport genes, the role of AdpC may be in signalling rather than transport. Our data expand the known ligands for TTT systems and identify a novel high-affinity binding protein for adipate, an important industrial chemical intermediate and food additive.

Databases
Protein structure co-ordinates are available in the PDB under the accession numbers 5OEI and 5OKU.

Introduction

Unlike the simple single-protein uptake systems such as those in the major-facilitator superfamily (MFS) [1], bacterial Substrate-Binding Protein (SBP)-dependent transporters can have substrate affinities in the nanomolar range, which make them excellent nutrient scavengers. These transporters contain, in addition to the transmembrane domains, periplasmic or extracytoplasmic SBPs, which bind to the substrate with high

Abbreviations
ABC, ATP-binding cassette; SBP, substrate-binding protein; TM, transmembrane; TRAP, tripartite ATP-independent periplasmic transporter; TTT, tripartite tricarboxylate transporter.

affinity, and deliver it to the transmembrane domains to be imported into the cell. Such transporters can be primary transporters, energised by ATP, such as the ATP-binding cassette (ABC) family [2,3], or secondary transporters, energised by ion electrochemical gradients such as the proton-motive force (pmf) and/or a sodium-ion gradient.

Two types of secondary SBP-dependant transporters are known; the TRipartite-ATP-Independent Periplasmic (TRAP) transporters [4] and the tripartite tricarboxylate transporters (TTT) [5]. The tripartite nature of these systems refers to the fact that transporters in both families are usually composed of three proteins or domains; a 10- or 12-transmembrane helix protein, responsible for transport of the substrate through the cytoplasmic membrane; a 4-transmembrane helix protein, the function of which remains unclear (although in TRAP systems it has been shown to be essential, [4]); and a periplasmic SBP [5]. Remarkably, despite the fact that TRAP and TTT systems are very similar in overall composition and organisation, they share no sequence similarity. The TRAP transporters, present in archaeal and bacterial species, were first characterised by Hamblin *et al.* [6] and Forward *et al.* [4] to transport C4-dicarboxylates. Diverse TRAP transporter SBPs and systems have since been extensively characterised and have been found to transport many different classes of compounds, but most often those containing carboxylate groups [7–9].

The TTT transporters, on the other hand, are mostly present in bacteria and rarely found in archaea [5], and their energy-coupling, physiological roles and substrates are far less well characterised than TRAP transporters. The *tctC* gene from *Salmonella enterica* encodes the prototype SBP for the family, originally shown by Sweet *et al.* [10] to bind citrate, but further members of the family were not characterised until Antoine *et al.* [11] found that this gene family was extensively overrepresented in *Bordetella pertussis*, with not only two sets of small and large transmembrane transport proteins but also with over 90 genes encoding TTT SBPs, not genetically linked to the transmembrane domains, thus being designated 'orphan' SBPs.

The functional significance of this surprising abundance of TTT SBP genes in some bacteria is unknown, but genome searches have revealed that orphan proteins in the TTT SBP family are commonly observed and found in many bacterial species, especially in soil bacteria, possibly reflecting their metabolic diversity, and that some of them are expressed at high levels, suggesting they might have important roles in the cell [11]. This is further confirmed by the fact that TTT SBP genes were the most abundant class of genes in metagenomics studies performed in the Amazon river, a complex and diverse environment [12]. Although it has been suggested that multiple paralog SBPs can function with a single transmembrane protein [13], they might be recruited by other transport families or their function might be related to signalling processes in the cell rather than transport [14]. The TTT SBP BctC, from *B. pertussis*, was found to be involved in a signalling cascade for citrate uptake [15], and it is believed that they can also be involved in stress responses and chemotaxis [16].

Only four TTT SBP structures have been elucidated so far, three of them from *B. pertussis*: BugD, bound to aspartate [17]; BugE, bound to glutamate [18]; Bug27, crystallised in an open conformation [16]; and TctC from *Polaromonas* sp., also in open conformation (access code 4X9T, unpublished data). The structures show that proteins in this family fold in the 'Venus flytrap' pattern typical of periplasmic solute-binding proteins, with two globular domains composed of β/α turns separated by a cleft, which folds around the substrate, providing a model for substrate binding which seems to apply for most proteins in this family. In recent years, a few TTT systems have been identified that transport diverse classes of substrates other than tricarboxylates or amino acids, such as nicotinic acid, nicotinamide and benzoate [16]; terephthalate and other aromatics [13]; 3-Sulfolactate [19]; and 3,3′-dithiodipropionic acid (DTDP) [20]. These studies raise the possibility that substrates for the TTT family might be far more diverse than previously thought.

One driver for the study of TTT systems is that they might be a new source of transporters for biotechnological processes. *Rhodopseudomonas palustris* is a purple nonsulphur bacterium that presents several biotechnological possibilities [21]. It has become a model organism for the study of the degradation of lignin-derived aromatic compounds, usually toxic even in low concentrations for most other bacteria and a limitation in the industrial degradation process [7,22,23]; for production of hydrogen gas in very high yields [24,25]; for generation of electricity in microbial fuel cells [26,27] and for bioremediation [28,29]. *Rhodopseudomonas palustris* has a complex and versatile metabolism, which allows photo- and chemotrophic growth and survival in a variety of environments. While most bacteria express around 6% of their protein-coding genes as transport-related proteins, *R. palustris* expresses more than 15%, which is likely to be related to its metabolic versatility and the different environments in which it can survive [21]. In this context, *R. palustris* may be a rich source for biotechnologically relevant transporters.

Genomic searches performed previously [11] showed that the TTT family in this bacterium is comprised of two complete tripartite systems (i.e. with each protein encoded by linked genes) plus five additional orphan SBPs. In this study, we show that one of these orphan SBPs (RPA4515), here designated AdpC, binds medium chain dicarboxylates, ranging from adipate (six carbons) to azelate (nine carbons), and that the *adpC* gene is upregulated in the presence of low adipate concentrations. By solving the crystal structure of AdpC with bound ligands, we identified conserved features for ligand coordination in the TTT family.

Results

Overproduction and purification of recombinant RPA4515

Protein BLAST searches in the *R. palustris* CGA009 genome using BugE from *B. pertussis* [18] as a protein query retrieved RPA4515 with 38% identity and 55% similarity, annotated as an uncharacterised protein containing a twin-arginine translocase (Tat) pathway signal peptide for periplasmic translocation. RPA4515 has 335 amino acid residues, of which 300 would be present in the 31.9-kDa mature protein after signal peptide cleavage. Genome context analysis of *rpa4515,* shown in Fig. 1A, shows no membrane components of any transport systems encoded nearby, but reveals the presence of some regulatory protein genes. *rpa4516*, immediately downstream of *adpC*, encodes a cyclic diguanylate phosphodiesterase (EAL-family) gene. On the opposite strand, *rpa4512*/*13* were shown by BLAST results to encode a two-component sensor-regulator system. Further bioinformatics investigations (not shown) predicted no transmembrane domains for the kinase component of this system, giving highest identities to HWE histidine kinases, cytoplasmic soluble kinases that respond to blue light [30].

Recombinant His-tagged RPA4515 was readily overproduced and purified in a single step by nickel affinity chromatography, as described in Materials and Methods. Elution from the nickel column was observed between 220 mM and 280 mM imidazole, and analysis of the eluate by SDS/PAGE, shown in Fig. 1B, shows the presence of the expected 35.6-kDa protein, equivalent to the His-tag and linker fused to the 31.9-kDa mature protein, as a single major band after fractionation. Some additional low abundance bands were observed, most likely breakdown products or copurifying minor contaminants. Mass spectrometry analysis of the purified sample gave an accurate mass of 35 549 Da for the RPA4515 protein (as predicted from the sequence) and peak integration of the chromatograms showed that the purity was ~96%. Size exclusion chromatography experiments, shown in Fig. 1C,D, demonstrated that the protein exists in solution as a monomer, as the elution peak was consistent with a protein of ~39 kDa. After discovery of adipate as a substrate (see below), a gel filtration experiment was performed in the presence of 1 mM adipate and the same elution pattern was observed, suggesting the protein remains as a monomer upon ligand binding.

Differential scanning fluorimetry shows that medium chain length dicarboxylates bind to RPA4515

RPA4515 was screened for ligand binding against a library of chemicals comprising 150 compounds of different classes (Table 1), using a differential scanning fluorescence assay in which the thermal denaturation curve of the protein will be slightly shifted due to the increased stabilisation provided by ligand binding. The raw denaturation curves in the presence of all substrates that improved protein thermal stabilisation are shown in Fig. 2A, the degree of thermal shift is plotted in Fig. 2B; the structure of the compounds used in these screens are shown in Fig. 2C. A shift of 9.7 ± 0.1 °C in the melting temperature was initially observed in the presence of the C6-dicarboxylate adipate (Fig. 2B). Expanding the screen to use compounds of similar structure, a shift of 5.2 ± 0.3 °C and 4.7 ± 0.1 °C was observed in the presence of 2-oxoadipate and *trans-trans*-muconate, respectively (Fig. 2B), but no thermal stabilisation was observed in the presence of hexanoate, 6-amino-hexanoate, 2-oxohexanoate, *cis-cis*-muconate, 6-amino-1-hexanol or hexamethylenediamine. The initial data thus suggested that two carboxylate groups and also a linear structure of the ligand are essential for coordination in the binding pocket of the protein and that neutral or positively charged functional groups were unable to provide stabilisation. In order to test whether longer and shorter chain lengths were also capable of providing thermal stabilisation, a new screen was performed with dicarboxylic acid salts from 4 to 9 carbons in length; pimelate (C7), suberate (C8) and azelate (C9) showed thermal stabilisations of 4 ± 0.1 °C, 6.3 ± 0.4 °C and 1.5 ± 0.1 °C, respectively (Fig. 2B), but no shift was observed for succinate, glutarate or sebacate, demonstrating that the protein could only accommodate dicarboxylates ranging from six to nine carbons in length in the binding pocket.

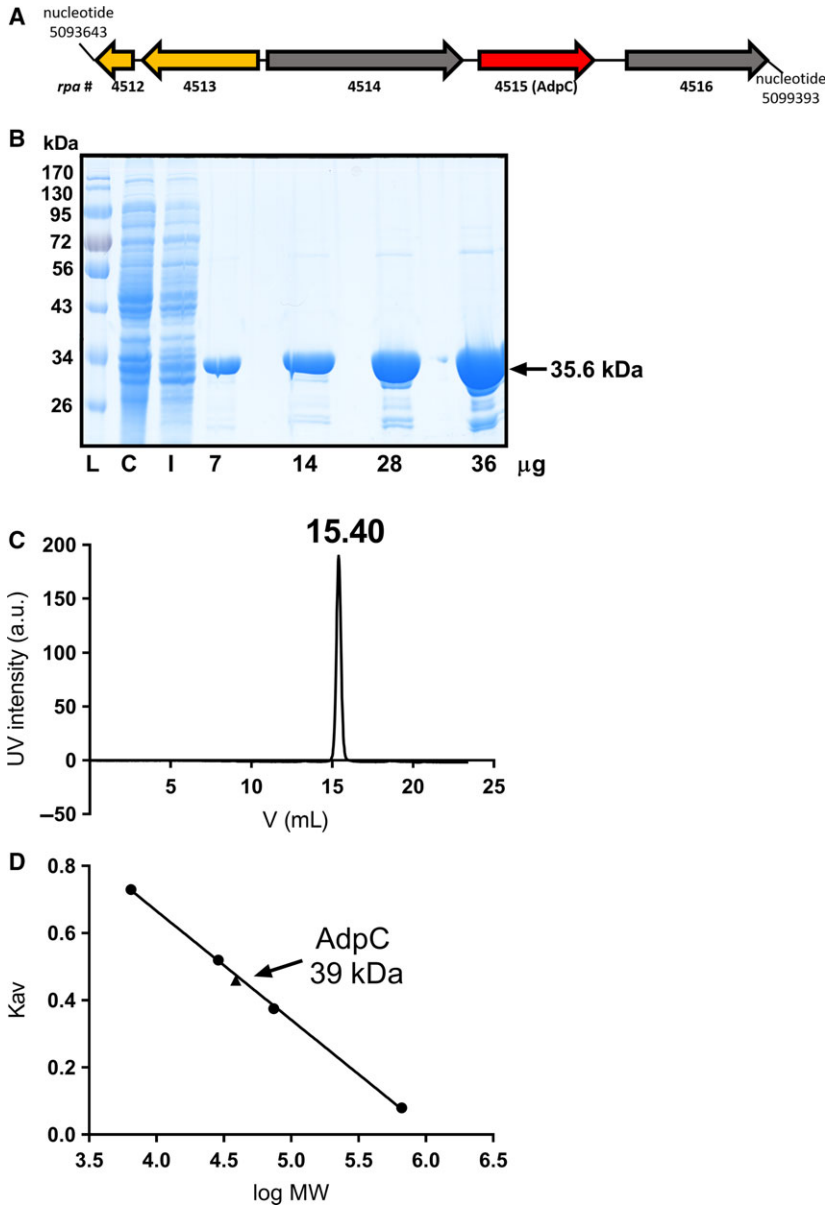

Fig. 1. Genomic context of *rpa4515* and purification of AdpC (RPA4515). (A) Genetic neighbourhood of *rpa4515* (*adpC*). *adpC* is just upstream of a cyclic diguanylate phosphodiesterase (EAL) gene, encoded by *rpa4516*. *rpa4514* encodes an alkaline phosphatase. *rpa4512/13*, on the opposite strand, were shown by BLAST results to encode a two-component sensor-regulator system, where RPA4513 is a member of the HWE histidine kinase family. (B) SDS/PAGE gel of AdpC purification. Samples were loaded in the following order: molecular weight size ladder (L); Control sample of *Escherichia coli* cell-fee extract (C); cell-free extract of induced strain (I), increasing protein volumes after purification and gel filtration in order to assess sample purity. The eluted protein corresponded well to the predicted size of AdpC of 35.6 kDa. (C) Size exclusion chromatography of AdpC. One milligram of protein in TF buffer (Tris 50 mM; NaCl 0.1 M; pH 7.4) was run on a 24 mL Superdex200 column at 0.5 mL·min^{-1} flow rate with the same buffer. A single peak was eluted at 15.40 mL V_e. (D) native AdpC molecular weight determination. Calibration curve for a Superdex 200 size exclusion chromatography column is shown, generated using thyroglobulin (669 kDa), conalbumin (75 kDa), carbonic anhydrase (29 kDa) and aprotinin (6.5 kDa). K_{av} was calculated using the formula $K_{av} = (V_e - V_o)/(V_t - V_o)$, as described in methods section. The elution of AdpC at 15.40 mL gives a K_{av} of 0.47, which is equivalent to a molecular weight of 39 kDa.

Isothermal Titration Calorimetry of RPA4515 reveals tight binding of adipate

The differential scanning fluorescence assay revealed important features regarding the substrate specificity for RPA4515. In order to obtain the thermodynamic parameters of binding and the affinity of the interactions, isothermal titration calorimetry experiments were performed. The corrected heat rates and normalised fits are shown in Fig. 3. The K_D values for the six tested substrates, shown in Table 2, reveal dissociation constants in the µM range and a ligand preference order similar to the degree of thermal shift in the differential scanning fluorescence assay, with adipate having the highest affinity with a K_D of ~0.5 µM. Considering this is the first report of an SBP with high affinity for this substrate and being consistent with the nomenclature for proteins belonging to the TTT transporters family, we propose RPA4515 to be designated as AdpC. The lower affinity of azelate binding, with a K_D 40 times higher than for adipate and considerably higher than all the remaining substrates, reinforces the conclusion that nine carbons are the upper limit supported for coordination in the binding pocket. Attempts to titrate AdpC with glutarate did not retrieve any heat change apart from the substrate dilution (data not shown), suggesting that six carbons is

Table 1. Library of compounds used for ligand screening in the differential scanning fluorescence assay.

Organic acids	Nitrogenated compounds	Aromatic acids	Sugars	Amino acids
2-Hydroxybutyrate	2,4,Dinitrophenol	4-Aminobenzoate	L-arabinose	L-alanine
2-Isopropylmalate	3,3,Diaminobenzidine	4-Chlorobenzoate	D-arabinose	L-arginine
Acetate	4-Acetamidophenol	4-Chlorocinnamate	D-erythrose	L-aspartate
Aconitate	4-Aminobutyrate	4-Hydroxybenzoate	D-fructose	L-cysteine
Adipate	4-Nitrocinnamate	4-Hydroxybutyrate	L-fucose	L-glutamate
Methyl malonate	*p*-Phenylene-diamine	4-Hydroxyphenylacetate	D-galactose	L-glutamine
Butyrate	4-Nitrophenylacetonitrile	4-Hydroxyphenyl-lactate	D-glucose	L-histidine
cis-cis-muconate	4-Nitrotoluene	4-Hydroxyphenylpropionate	D-lactose	L-isoleucine
Citrate	4-Phenylbutyronitrile	4-Hydroxyphenylpyruvate	D-lyxose	L-leucine
Crotonate	6-Amino-1-hexanol	4-Methoxycinnamate	D-maltose	L-methionine
Fumarate	Adenine	4-Phenylbutyrate	D-manitol	L-phenylalanine
Gluconate	Adipamide	5-Phenylvalerate	L-rhamnose	L-proline
Glutaconate	Adiponitrile	Benzoate	D-ribose	L-threonine
Glutarate	Aminolevulinate	Caffeate	D-ribulose	L-tryptophan
Itaconate	Benzamide	Cinnamate	D-sucrose	L-valine
Lactate	Fumaronitrile	Coumarate	D-tylose	
Levulinate	Glutaronitrile	Ferrulate	D-xylose	
Malate	Hexamethylenediamine	Gallate		
Malonate	Indole-3-acetamide	Phenylacetate		
Mesaconate	Mandelamide	Phenylpyruvate		
Methyl-2-oxobutanoate	Metronidazole	Protocatechuate		
Methyl heptanoate	Orotate	Syringate		
Methyl hexanoate	4-Nitrophenylsulphate	Vanillate	**Others**	**Vitamins**
Propanoate	4-Nitrophenol	Vanillin	Adonitol	Ascorbate
Methyl octanoate	Sulphanilamide		Barbitone	Biotin
n-Valerate	*trans*-nitrostyrene		EDTA	Nicotinate
Oxalate	Triethanolamine		Fucitol	Pyroxidine
Oxaloacetate			TAPS	Thiamine
Succinate			TES	D-panthotheate
Tartrate			Glyceraldehyde	Nicotinamide
α-Ketoglutarate			Taurine	
β-Hydroxypyruvate			Ectoine	

the shortest chain length possible for an efficient interaction. The data also revealed that while binding of dicarboxylates comprising eight and nine carbons produces an exothermic reaction, those of six and seven carbons generate an endothermic effect.

adpC expression is induced at low concentrations of substrate

In order to complement the information provided by the binding assays and genomic context of the *adpC* gene, RT-PCR experiments to study *adpC* expression were performed in the absence or presence of dicarboxylates during growth of *R. palustris* in minimal media. In a first approach, expression of *adpC* was evaluated in the presence of 10 mM of various substrates. As shown in Fig. 4A, instead of an expected induction, a fivefold lower expression was observed in the presence of adipate and pimelate, and a 10-fold lower expression in the presence of suberate and azelate. Given that the TTT transporters are high-affinity transporters, it was speculated whether *adpC* might be more highly expressed under much lower (and more environmentally relevant) substrate concentrations. An expression analysis was therefore performed using different concentrations of adipate. As shown in Fig. 4B, a 5.8-fold increase in *adpC* expression was observed in the presence of 1 μM adipate compared to the absence of this compound in the growth medium. As the adipate concentration was increased to 10 μM, 100 μM and 1000 μM, however, *adpC* expression decreased, showing a 4.3-, 4.1- and 2.4-fold expression when compared to the control. These data suggest *adpC* expression is carefully modulated by the concentration of extracellular adipate. In order to determine whether *R. palustris* could actually use adipate as the sole carbon source, a growth experiment was performed under phototrophic conditions in minimal media in the

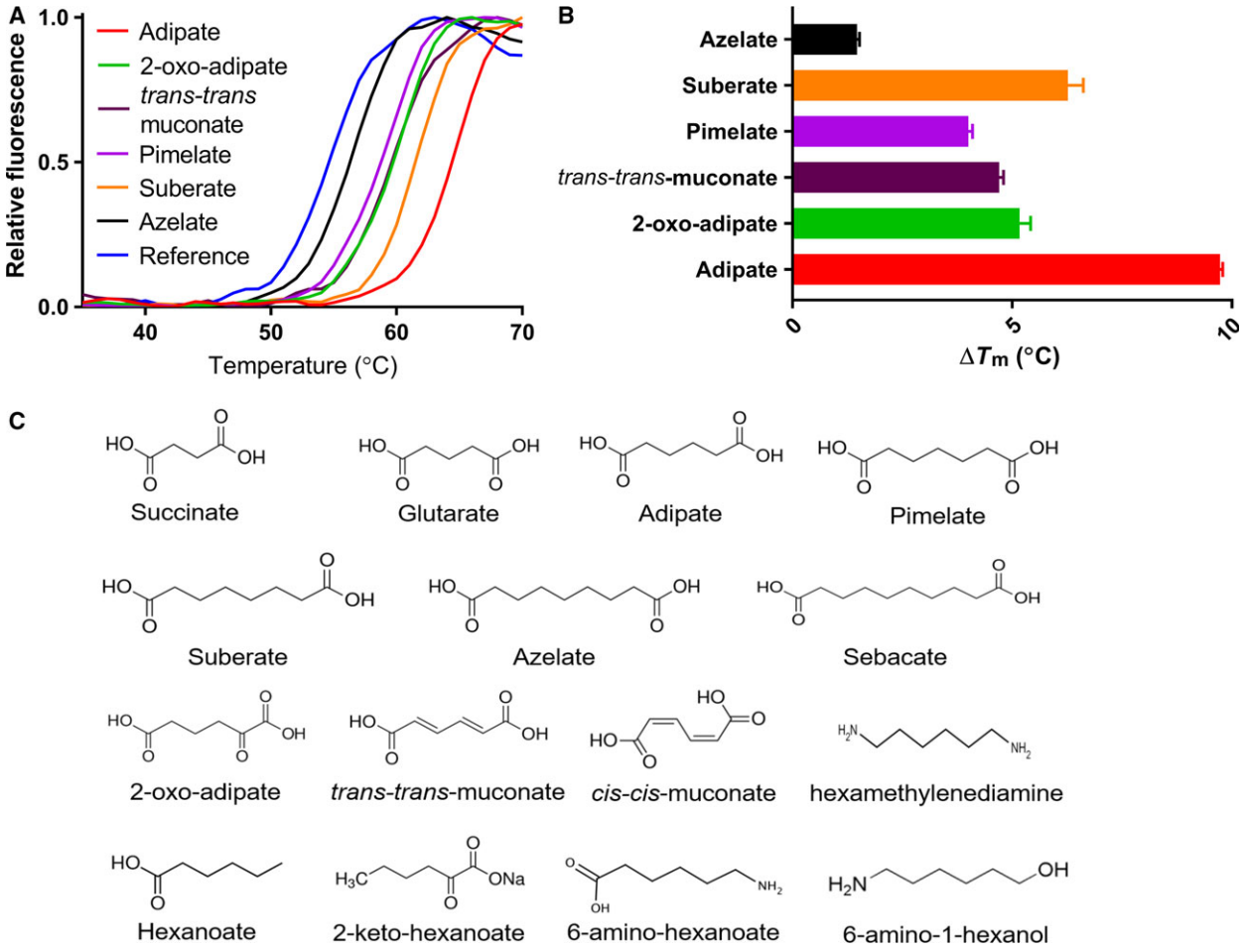

Fig. 2. Differential scanning fluorimetry of AdpC. Each well of a 96-well optical plate contained 50 μL TF Buffer (Tris-HCl 50 mM pH 7.4 NaCl 0.1 M), 5 μM of AdpC, 60 μM of ligand and 1x SYPRO orange™ dye, and fluorescence was monitored from 25 °C to 94 °C. Experiments were performed in triplicate. (A) Raw data of the denaturation curves in the presence of different dicarboxylates. (B) Histograms of the ΔT_m (°C) of AdpC in the presence of different dicarboxylates. The data represent the mean and standard deviation of the three replicates. The following shifts were observed in the presence of the respective ligands: Adipate (red): 9.7 ± 0.1 °C; 2-oxoadipate (green): 5.2 ± 0.3 °C; *trans-trans*-muconate (purple): 4.7 ± 0.1 °C; pimelate (pink): 4 ± 0.1 °C; suberate (orange): 6.3 ± 0.4 °C; azelate (black): 1.5 ± 0.1 °C. (C) Structures of compounds similar to adipate were tested in the analysis.

presence of 10 mM adipate. As shown in Fig. 4C, both the growth rate and cell yield (final OD_{660}) were significantly higher with adipate compared to growth with the same concentration of the C4-dicarboxylate succinate. These findings, however, do not show growth on adipate is mediated by AdpC, and our attempts to construct an *adpC* knockout mutant proved unsuccessful.

High-resolution crystal structures of AdpC with adipate and 2-oxoadipate show conserved coordination mechanisms in the TTT family

Purified AdpC was readily crystallised in the presence of 2 mM ligands and essentially isomorphous structures were obtained with bound adipate and 2-oxoadipate. The AdpC crystal structure was solved using molecular replacement and the overall structure of the 2-oxoadipate bound protein is shown in Fig. 5 and described below. Table 3 summarises the data collection and refinement statistics for AdpC bound with either adipate or 2-oxoadipate.

AdpC is a monomer comprising of 300 amino-acid residues, formed by residues 34–334, where residues 36–334 are included in the 1.8Å structure. It contains nine β-strands (β1–β9), nine α-helices (α1–α9) and three small helices, composed of just three amino acid residues. β-strand 4 is disrupted by Pro136, and for this reason is renumbered β4a and β4b. This secondary structure is organised in the typical type II SBP shape

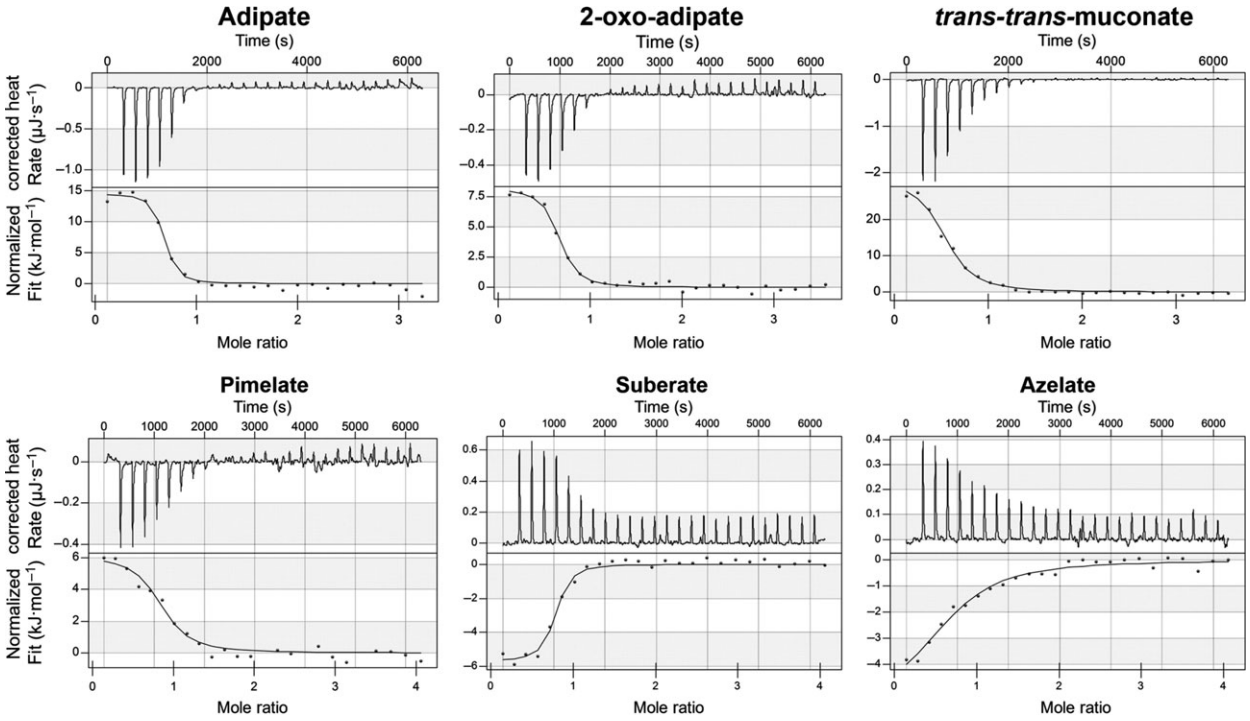

Fig. 3. Isothermal titration calorimetry of AdpC with dicarboxylates ranging from six to nine carbons chain length. Experiments were performed with 300 μL of 100 μM protein in TF Buffer (Tris-HCl 50 mM pH 7.4 NaCl 0.1 M), with 25 × 2 μL ligand injections with 4 min between each injection. Corrected heat rates are shown in the top panels and a normalised fit in the bottom panels for each respective ligand. Negative heat rates imply endothermic interactions, whereas positive heat rates imply exothermic interactions.

Table 2. Summary of thermodynamic properties of dicarboxylate binding to AdpC (RPA4515). Isothermal titration calorimetry was used to determine binding properties as described in Materials and Methods**.** The tightest binding affinity is observed with adipate, while the titration with azelate retrieved a K_D 40 times higher. Attempts to titrate the protein with glutarate (a C5 dicarboxylate) showed no heat change. C6 and C7 dicarboxylates generated an endothermic effect; chains of eight and nine carbons in length, on the other hand, caused an exothermic effect upon binding. (*n*) represents the protein:ligand stoichiometry. The value in parentheses following the ligand name is the carbon chain length.

Ligand	K_D (μM)	ΔH (kJ·mol^{-1})	ΔS (J·mol^{-1}·K^{-1})	ΔG (kJ·mol^{-1})	*n*
Adipate (C6)	0.55 ± 0.49	14.51 ± 0.95	167.70	−36.32	0.63 ± 0.02
2-Oxoadipate (C6)	1.60 ± 0.67	8.16 ± 0.39	137.90	−33.63	0.61 ± 0.02
trans-trans-muconate (C6)	4.66 ± 1.45	30.91 ± 2.05	204.00	−30.94	0.52 ± 0.02
Pimelate (C7)	3.60 ± 2.78	6.14 ± 0.69	124.50	−31.59	0.82 ± 0.07
Suberate (C8)	0.93 ± 0.60	(−)5.70 ± 0.33	96.70	−35.01	0.73 ± 0.03
Azelate (C9)	20.02 ± 10.46	(−)5.50 ± 1.34	71.80	−27.27	0.69 ± 0.12

[31], a venus flytrap-like conformation, with two globular domains of β-sheets of five strands each surrounded by α-helices, separated by a cleft, where the substrate is coordinated (Fig. 5). The hinge connecting the two domains is composed of β-strand 4 and a loop between h3 and β9, characterising AdpC in the Cluster E-II for SBP classification according to Scheepers *et al.* [31]. Domain 1 is comprised of residues 36–135 and 261–334, and domain 2 is comprised of residues 136–260 from the N to C terminals. The β-sheets are topologically formed by β2-β1-β3-β9-β4a in domain 1 and β6-β5-β7-β4b-β8 in domain 2. Unlike the other members of the TTT SBP family structurally characterised so far [11,17,18], AdpC contains no cysteines and therefore, no disulphide bridges.

The full substrate coordination of 2-oxoadipate is shown in the LigPlot diagram in Fig. 6A and in comparison to aspartate and glutamate binding in Bug D and BugE, respectively, in Fig. 6B. The structure with 2-oxoadipate shows a conserved mechanism for

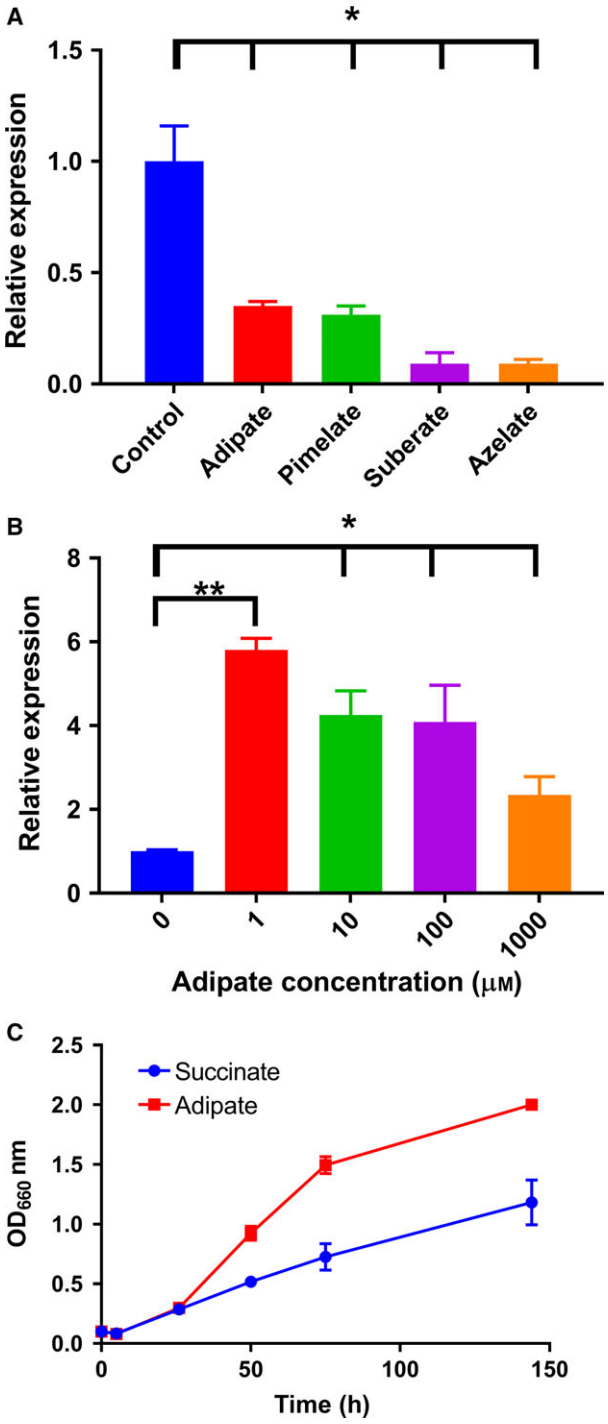

Fig. 4. The RT-PCR experiments showing the expression levels of the *adpC* gene in *Rhodopseudomonas palustris* after growth with various dicarboxylates. (A) Expression levels after growth in the presence of 10 mM dicarboxylates. The data show the mean and standard deviation of three replicates. A fivefold repression is observed in the presence of adipate and pimelate, and a 10-fold repression in the presence of suberate and azelate. (B) Expression levels under induction with different adipate concentrations. A 5.8-fold increase is observed in the presence of 1 μM adipate, with a 4.3-, 4.1- and 2.4-fold expression under induction with 10 μM, 100 μM and 1000 μM adipate. Statistics were performed with Student's *t*-test (*$P < 0.05$ and **, $P < 0.01$). (C) Growth curves in the presence of 10 mM succinate or adipate as the sole carbon source. Data shown represent the mean and standard deviation of three replicate cultures. Both the growth rate and final OD_{660} were found to be higher in the presence of adipate when compared to succinate.

coordination of the proximal carboxylic group between AdpC, BugD and BugE, with participation of a β-turn-α segment between β1 and α1 and between β6 and α5 forming a 'pincer' structure, as described previously [16], with a Thr54 side chain and Thr49 and Ser198 main chains participating in this process. Two water molecules play a major role in coordination, bridging several hydrogen bonds with the main chain of three residues present in the pincer structure, Thr49, Asp55 and Ala199 (Fig 6A,B). Although it is the main chain, and not the side chains of the residues, which are responsible for most interactions, these residues are well conserved among the TTT family [11], with the exception of Thr49. Hydrophobic interactions also play a role in ligand coordination. Phe48 is very well conserved among the TTT SPBs [11] and seems to act as a hydrophobic platform for the aliphatic regions of the substrates characterised so far. Gly52 and Gly197 are also conserved, and seem to help organise the structure around the carboxylic group of the substrate. The carbonyl group in 2-oxoadipate is bound by the side chain of two serines, Ser169 and Ser198 (Fig. 6A), conserved in BugE but not BugD. The distal carboxylate group on the substrate is coordinated by hydrogen bonds with the side chain and main chain of Thr174 and the side chain of Ser261 and Asn216, and interestingly, by two additional water molecules which bridge hydrogen bonds with Gly106, Ser173, Gly175 and Asp215 (Fig. 6A,B). The residues and water molecules coordinating this distal carboxyl group are much less conserved, where only Asp215 is present in the three proteins. Gly106 is present in BugD and Asn216 is also found in BugE and is present as a conservative change as a Gln in BugD, but absent in Bug27. Interestingly, there are no positively charged residues in the binding pocket to counteract the negative charge of the substrate carboxylate groups, as is well described for arginine residues in TRAP transporter SBPs [9].

Discussion

In this study, we characterised the first SBP belonging to the TTT family in *R. palustris* through genetic and

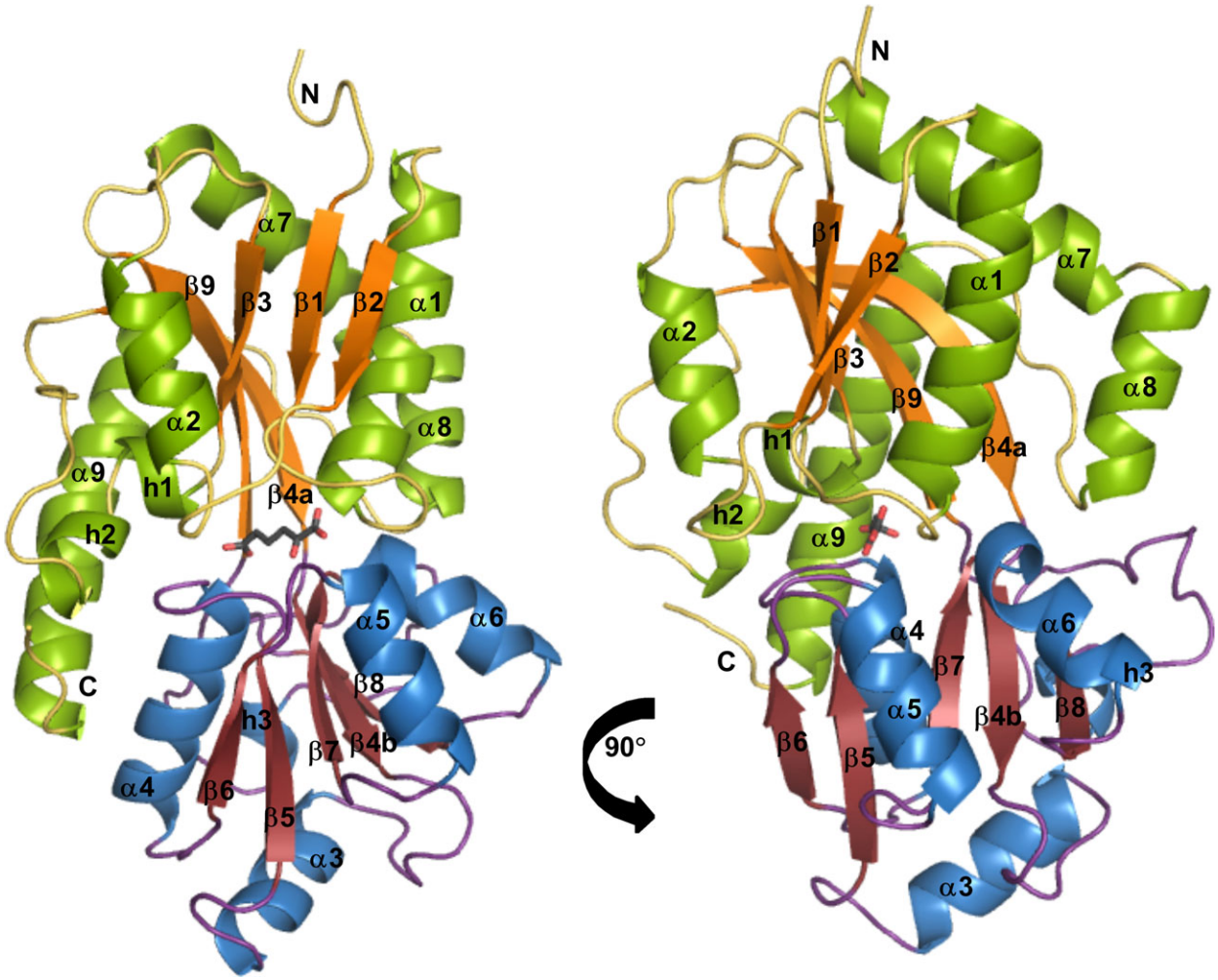

Fig. 5. A 1.8-Å structure of AdpC, with 2-oxoadipate coordinated in the binding pocket. The overall structure reveals a venus flytrap-like conformation, comprised of two globular domains separated by a cleft, connected by β4 and a loop, characterising the protein as belonging to cluster E-II in the SBP classification. Each domain is constituted of a 5-strand β-sheet surrounded by α-helices. The β-sheets are topologically formed by β2-β1-β3-β9-β4a in domain 1 and β6-β5-β7-β4b-β8 in domain 2. Substrate coordination occurs in part due to a 'pincer' structure, formed by two β-loops between β1 and α1 and β6 and α5.

transcriptional analysis, ligand binding properties and 3D structure determination with two bound ligands. Genome analysis did not show any membrane component of a transport system coencoded with the *rpa4515* SBP gene, characterising it as encoding an orphan SBP. Such genes are frequently found in higher numbers than the membrane components, especially in β-proteobacteria, where over a hundred SBP genes can sometimes be found in contrast to perhaps a few complete systems, as shown by Antoine *et al.* [11] for *Bordetella*. As a poorly characterised family, it is as yet unclear whether these SBPs interact in a promiscuous manner with the much smaller number of transmembrane proteins or participate in signal transduction pathways, but the participation of other types of SBPs in signalling processes is known and has been reviewed by Piepenbreier *et al.* [14]. In the study performed by Antoine *et al.* [15], it was shown that the citrate uptake SBP BctC from *B. pertussis* is not only involved in citrate transport but also plays a role in gene regulation, interacting with the two-component system BctDE in order to modulate expression levels of the *bctCBA* system itself. Participation of TTT SBPs in signalling processes were further suggested for *Bacillus subtillis* and *Salmonella typimurium* [5,32]. Although a two-component system is directly upstream of *adpC* on the opposite strand, bioinformatics analysis showed that the sensor domain in RPA4513, contains no transmembrane regions and encodes for a histidine kinase of the HWE family, which are soluble

Table 3. Data collection and refinement statistics for AdpC. Values in parentheses are for the outer resolution shell.

Data collection	AdpC + 2-oxoadipate	AdpC + adipate
Wavelength (Å)	0.92819	0.97628
Resolution range (Å)	59.0–1.78 (1.81–1.78)	41.65–2.07 (2.11–2.07)
Space group	C 2 2 2_1	C 2 2 2_1
Cell dimensions		
(a,b,c) (Å)	47.3, 168.5, 82.5	46.9, 166.6, 82.2
(α,β,γ) (°)	90, 90, 90	90, 90, 90
Total reflections	234 875 (11 301)	189 647 (9713)
Unique reflections	32 190 (1594)	20 144 (1010)
Multiplicity	7.3 (7.1)	9.4 (9.6)
Completeness (%)	100 (99.6)	100 (100)
Mean I/σI	10.5 (2.1)	5.0 (2.3)
Rmerge	0.15 (0.95)	0.63 (2.59)
Rmeas	0.17 (1.03)	0.67 (2.76)
Rpim	0.08 (0.53)	0.21 (0.9)
CC half	0.997 (0.572)	0.977 (0.490)
Refinement		
Resolution range (Å)	58.9–1.78	41.65–2.07
No. of reflections used	30 497	19 116
No. of reflections in R_{free} set	1671	998
R_{cryst}/R_{free}	0.181/0.217	0.207/0.254
No. of protein protomers	1	1
No. of protein atoms	2242	2237
No. of ligand atoms	11	10
No. of water molecules	111	101
R.m.s.d., bonds (Å)	0.021	0.017
R.m.s.d., angles (°)	1.961	1.893
Mean *B* values (Å^2)		
Protein	19	22
Ligand	18	20
Water	23	25
Ramachandran plot (%)		
Favoured	98.3	98.0
Allowed	1.7	2.0
Disallowed	0.0	0.0
PDB code	5OEI	5OKU

in the cytoplasm and are known to respond to light stimuli [30]. Thus, it seems unlikely that AdpC interacts with this kinase but further investigations are needed to determinate if there is a membrane transport protein that AdpC interacts with and indeed the precise physiological function of AdpC.

The differential scanning fluorescence assay showed evidence for binding of dicarboxylates ranging from six to nine carbons in length. The binding of *trans-trans*-muconate but not *cis-cis*-muconate revealed the need for a linear substrate structure. Furthermore, the absence of binding of 1-hexanol suggested that both carboxylate groups were necessary for binding, and that replacing one or two of these groups with polar or positively charged groups did not result in binding, as shown by the lack of interaction with 6-amino-1-hexanol and hexamethylenediamine. The few TTT SBPs characterised so far seem to bind to substrates containing at least two or more carboxylate groups, as observed with the citrate binding TctC [33], the terephthalate-binding protein TphC [13], the disulphide 3,3′-dithiodipropionic acid (DTDP)-binding protein TctC [20] and the sulpholactate-binding protein SlcH [19], where a sulphate group gives a similar polarity to the molecule as a carboxylate group would, raising the question as to whether this would be a prevalent characteristic of this family. So far the only exception to this pattern is the Bug27 protein, which was shown to bind to nicotinate [16].

Isothermal titration calorimetry assays showed μM range dissociation constants for a range of dicarboxylates. Although the highest affinity was observed with adipate, with a K_D of ~0.5 μM, tight binding was maintained with pimelate and suberate, but azelate showed a 40-times lower affinity, reinforcing the notion that nine carbons is the upper limit that the binding pocket can accommodate. Substrates of six and seven carbons in length showed an endothermic effect, in a binding event likely to be driven by the entropy generated from displacement of solvent molecules and hydrogen bonds in the binding pocket, using the energy of positive enthalpy change derived from the ligand molecule kinetic and hydrophobic interactions [7]. The exothermic binding of the longer chain compounds, on the other hand, is driven to a minor extent by this effect, and more by the negative enthalpy deriving from the formation of hydrogen bonds, suggesting hydrogen bonds and solvent displacement have different weights in each case [34].

Structures of AdpC were successfully obtained with either adipate or 2-oxoadipate bound. The 1.8 Å structure with 2-oxoadipate in the binding pocket reinforced the ubiquity in the TTT family of a pincer structure shown to coordinate the ligand's proximal carboxylate group formed by β-loops between β1 and α1 and β6 and α5, making hydrogen bonds bridged by two water molecules in very conserved positions [16–18]. Further structures might reveal whether this is a feature common to the whole family, as suggested by sequence alignments [11]. Our structures also revealed an absence of positively charged residues in the binding pocket to counterbalance the negative charge of the substrate, and although it might be expected that substrates containing amino groups would better suit the pocket due to its small overall net negative charge caused by the presence of Asp215, our data strongly

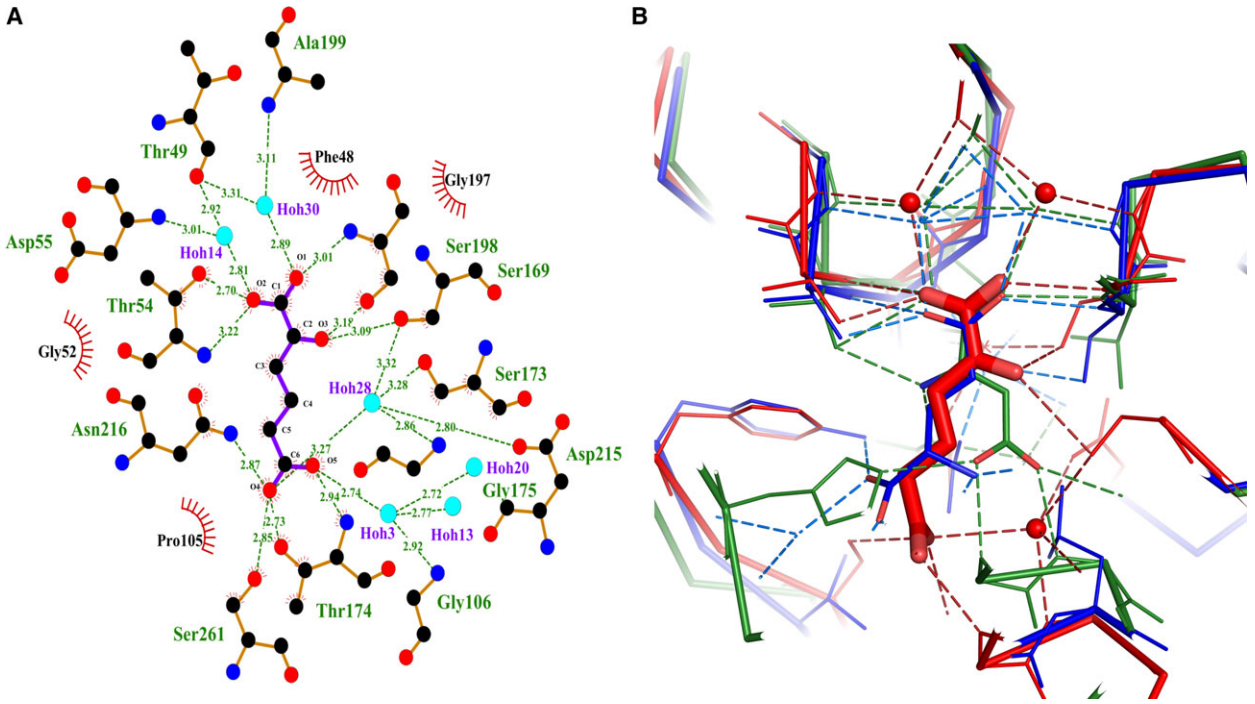

Fig. 6. (A) Ligplot showing coordination of 2-oxoadipate in AdpC-binding pocket. Water molecules play an important role in bridging hydrogen bonds with the main chain in surrounding residues of the 'pincer' structure. Interestingly, no positively charged residues are found to counteract the substrate negative charge. The keto oxygen on the 2-oxoadipate ligand is coordinated by the side chains of Ser169 and Ser198; (B) Comparison between substrate coordination for 2-oxoadipate in AdpC (red), aspartate in BugD (green) and glutamate in BugE (blue). The two water molecules coordinated by the 'pincer' structure are conserved in position (red spheres at the top of the figure). However, most of the other residues involved in the water coordination are not conserved among the three proteins. The coordination of the second carboxylic group of each substrate is not only much less conserved but also seems to involve at least one water molecule. Dashed lines represent hydrogen bonds between the substrate and the binding pocket.

suggest otherwise. This repulsive negative charge might, however, be compensated somewhat by the dipoles of helices α1 and α4 and possibly to some extent α5 which are directed towards the carboxylate groups of the substrate. In addition, the solvation provided by the presence of water molecules may dissipate the charge, as shown previously for TRAP transporters [35]. Given the space restrictions created in the binding pocket of AdpC by h1 and the N-terminal region of α4 in one end and the 'pincer' structure in the other, it is unclear how this protein is able to accommodate substrates containing up to nine carbons in length and this will require further structural studies. Overall, using water molecules as part of the binding mechanism might give more flexibility in the types of ligands that can be bound in a given TTT SBP, exemplified for example by the Bug27 protein (nicotinate, nicotinamide, citrate, benzoate, quinaldic acid) and TphC (terephthalate, protocatechuate) [13,16]. Identification of additional ligands for uncharacterised TTT SBPs should reveal the extent of this versatility in this family of substrate-binding proteins.

The study of *adpC* expression by RT-PCR showed that the presence of 10 mM of any of the ligands tested resulted in a 5–10-fold lower expression of the *adpC* gene. Further investigation showed that a 5.8-fold increase in expression was observed in the presence of 1 μM adipate, and that this increased expression became progressively smaller as the adipate concentration was increased. If AdpC is indeed involved in transport rather than signalling, it is possible the high-affinity properties of the TTT transporter might be only required at low-substrate concentrations. Uptake in higher substrate concentrations might be performed by other types of secondary transporters, such as the ones from the MFS family [1]. On the other hand, if AdpC acts as a signal transduction protein, its synthesis may need to be exquisitely sensitive to low environmental concentrations of adipate. In either case, the adipate-dependent expression of *adpC* suggests that a fine-tuned regulatory system exists to modulate AdpC production. Harrison and Harwood [36] demonstrated the ability of *R. palustris* to grow in minimal media using dicarboxylic acids from seven to nine carbons in

length as the sole carbon sources, and our observations show it can also grow well in the presence of the C6 dicarboxylate adipate. In nature, adipate is a known intermediate of a cyclohexanol degradation pathway, which is then metabolised by β-oxidation [37]. In *R. palustris*, anaerobic aromatic acid degradation generates cyclohexane carboxylate and pimelate as intermediates [36,38]. Considering that the affinity constants of AdpC for adipate, 2-oxoadipate, pimelate and suberate are rather similar, it might be that the physiological role of AdpC involves binding of several of these related dicarboxylates. 3-oxoadipate, for example, is a central intermediate in the degradation of protocatechuate and catechol, via a pathway widespread among soil bacteria, including *R. palustris* [21,39]. Moreover, pimelate is a key intermediate in the degradation of lignin-derived aromatic compounds in the benzoate catabolic pathway after the ring is cleaved [23] and is further degraded through β-oxidation in *R. palustris* by proteins encoded by the *pimFABCDE* operon, which would also be responsible for suberate degradation [36]. Encoded adjacent to this operon are ABC SBPs which were shown to bind to a range of dicarboxylates and fatty acid analogs, ranging from 7 to 14 carbons in length, but interestingly not to adipate [40]. As a precursor, pimelate production is needed in the cell for biotin synthesis [41,42]. Dicarboxylic acids are a key class of molecules for photosynthetic bacteria, representing for many of them the best carbon sources for growth [43]. Thus, the presence of high-affinity binding proteins in *R. palustris* for these compounds might reflect the lignin-rich environments in which this bacterium is found, potentially helping to regulate the chemotaxis, uptake and metabolism of these aliphatic degradation products of aromatic compounds [44,45].

Our identification of a high-affinity, adipate-binding protein is also relevant in the context of biotechnology. Adipate is one of the most frequently used carboxylates industrially, with a global production of ~2.6 million ton/year [46]. Its main use is as a precursor for nylon 6-6 fibres, but it is also used in polyurethane production and as a food additive in foods that require acidity regulation [47]. Conventionally produced from benzene chemical transformations, there is intense research aiming to produce adipate from renewable sources such as lignin, lipids and TCA cycle intermediates, using both natural and synthetic pathways, as reviewed by Kruyer and Peralta-Yahya [48]. An adipate biosensor would be useful in measuring adipate concentrations and optimising yields from such processes. AdpC could be employed in this way, by detecting the conformational change upon ligand binding using a variety of biophysical techniques, as has been extensively demonstrated for the maltose-binding protein, reviewed by Medintz *et al*. [49,50].

In conclusion, the present study has described the first SBP from the TTT family to bind with high affinity to medium chain length dicarboxylates, showing that this family is involved in the metabolism of a wider variety of substrates than those recognised so far. Moreover, the binding pocket of AdpC shows flexibility, binding to substrates from six to nine carbons in length. We also show that AdpC expression is regulated by low substrate concentrations, suggesting that it may play a role in regulatory switching for different substrate availability. Taken together, these data help to further characterise the mechanisms and physiological role of TTT systems.

Materials and methods

Media and strains

Rhodopseudomonas palustris CGA009 (ATCC® BAA-98™) was used in this study [51]. Growth in rich media was in PYE, composed of 5 $g{\cdot}L^{-1}$ each of peptone, yeast extract and succinate. RCV was used as minimal media, using 40 mM succinate as carbon source [52]. Liquid growth was performed at 30 °C anaerobically in the light, with two 9W warm white LED bulbs positioned at 20 cm from the cultures, while growth on solid agar media was performed aerobically in the dark. All *Escherichia coli* strains were grown in LB media at 37 °C. When used, antibiotic concentrations for *R. palustris* and *E. coli* were respectively: Chloramphenicol 20 $\mu g{\cdot}mL^{-1}$ and 34 $\mu g{\cdot}mL^{-1}$; Carbenicillin (not used) and 50 $\mu g{\cdot}mL^{-1}$; gentamycin 100 $\mu g{\cdot}mL^{-1}$ and 20 $\mu g{\cdot}mL^{-1}$.

Protein overproduction and purification

CGA009 genomic DNA was extracted with the 'GenElute™ Bacterial genomic DNA kit' (Sigma, Poole, UK). The *rpa4515* gene was PCR amplified using the primers rpa4515_F (5′-ATAGAGCTCAGACTGGCCGACCCGG-3′, SacI site underlined) and rpa4515_R (5′-TTAAAGCTT CTACAGCGGCGCAATTCC-3′, HindIII site underlined), which excludes the protein N-terminal signal sequence comprising of the first 35 amino acids. The amplified DNA was cloned into the pBAD-HisB vector (Invitrogen, Carlsbad, CA, USA) through SacI and HindIII restriction sites to add an N-terminal His_{6x} tag to the protein. The resulting pBAD-4515 plasmid was transformed into *E. coli* DH5α and subsequently into *E. coli* Top10(DE3) (Invitrogen) for overexpression. *rpa4515* was overexpressed under the control of the $_{ara}$BAD promoter contained in pBAD-4515. *E. coli* TOP10 (DE3)(pBAD-4515) cells were grown to OD_{600} 0.6 under

carbenicilin selection and then induced with 0.2% (w/v) of L-arabinose for 24 h at 37 °C before harvesting by centrifugation (16 000 ***g***, 10 min). Harvested cells were resuspended in 20 mL of Binding Buffer (20 mM sodium phosphate buffer pH 7.4, 500 mM sodium chloride, 20 mM imidazole) and lysed by sonication (MSE soniprep 5 × 20 s at 16 microns amplitude with cooling by ice). The resulting cell-free extracts (CFE) were fractionated on a 5 mL HisTrapTM HP crude column (GE Healthcare, Little Chalfont, UK), with the recombinant protein eluted using a 0–500 mM imidazole gradient (Elution buffer; 20 mM sodium phosphate buffer pH 7.4, 500 mM sodium chloride, 500 mM imidazole). Size exclusion chromatography of RPA4515 was performed using a 24-mL Superdex200 column at 0.5 mL·min^{-1} flow rate in an AKTA protein purification system, using TF Buffer (Tris 50 mM; NaCl 0.1 M pH 7.4). A calibration curve was generated using thyroglobulin (669 kDa), conalbumin (75 kDa), carbonic anhydrase (29 kDa) and aprotinin (6.5 kDa), and K_{av} was calculated using the formula $K_{av} = (V_e - V_o)/(V_t - V_o)$, where V_e is the elution volume for a given protein, V_o is the void volume and V_t is the total volume of the column.

RT-PCR of *rpa4515*

Rhodopseudomonas palustris was grown phototrophically in RCV minimal media [52], using 10 mM succinate as the sole carbon source, until an OD_{660} of 0.7 was reached. In one set of experiments, carboxylic acids (pyruvate, adipate, pimelate, suberate or azelate) were then added to replicate cultures to a final concentration of 10 mM. Adipate was added to separate replicate cultures to 1 μM, 10 μM or 1000 μM final concentrations. Cells were then incubated for 2 h before harvesting by centrifugation at 16 000 ***g*** for 10 min. RNA was then extracted as described by Guccione *et al.* [53]. Primers for RT-PCR were designed using PRIMER3 software (http://primer3.ut.ee/), aiming for 20-bp length, with a melting temperature of 58 °C and amplifying a central region of the gene between 200 and 400 bp in length, at least 100 bp after the mRNA start position. For *rpa4515*, primers used were AdpC_FW (5′-CGGCCAGTCCTTTGTCAT-3′) and Adpc_rev (5′-GATCGTAGCCCATGGTCC-3′), and for the housekeeping gene *rpoD* primers were RpoD_FW (5′-CGACTTCCTGCGCAACTATC-3′) and RpoD_rev (5′-GGTTGGTGTACTTCTTGGCG-3′). Each reaction was carried out in a 25 μl volume in a MicroAmp® 96-well optical reaction plate (ABI prism). Reactions were performed using the Brilliant III Ultra-fast SYBR Green RT-PCR kit (Agilent, Santa Clara, CA, USA), according to the manufacturer's instructions. Each reaction using RNA was repeated in biological triplicate, each with a technical triplicate; reactions using genomic DNA for the standard curve, were replicated in duplicate. PCR amplification was carried out in a Stratagene MX3005p thermal cycler (Agilent) according to the manufacturer's instructions. Data were collected with the associated MXPRO QPCR software (Agilent). A standard curve for each gene was generated using a series of *R. palustris* genomic DNA dilutions. Gene expression between cultures was calculated as relative to *rpoD* expression. The data were analysed as described previously [53].

Differential scanning fluorescence assay

Thermal stability assays were performed according to Vedadi *et al.* [54], in a Mx3005P RT-PCR machine (Stratagene, La Jolla, CA, USA), in a 96-well plate format, each well containing 50 μL of 5 μM protein, 60 μM of ligand, 1x SYPRO Orange Dye (Invitrogen) final concentrations in TF Buffer (Tris-HCl 50 mM pH 7.4 NaCl 0.1 M). Initial screening was performed against a library comprising of 90 compounds from different classes (aliphatic, amino acids, dicarboxylic acids, aromatics and vitamins). The composition of this library is given in Table 1.

Isothermal titration calorimetry

Isothermal titration calorimetry analyses were carried out in a Nano ITC calorimeter (TA instruments, New Castle, DE, USA). RPA4515 was concentrated to 100 μM and dialysed against TF Buffer (Tris-HCl 50 mM pH 7.4; NaCl 0.1 M). The dialysis buffer was used to prepare stock substrate solutions at 1 mM. The reaction cell contained 300 μL of 100 μM protein. Titrations were carried out at 30 °C by 25 × 2 μL injections, with a 4-min interval between injections. After optimising the baseline values and discounting the values of ligand dilution, integrations, fitting to an independent-binding model, K_D determination and statistics were performed using NANOANALYZE software (TA instruments).

Protein crystallisation, data collection and structure determination

Protein crystallisation was done by concentrating the protein in TF buffer to 10 mg·mL^{-1} with addition of 2 mM of ligand and testing in sitting drop experiments with a range of different commercial screens (Molecular Dimensions Ltd, Newmarket, UK). Screening conditions that produced crystals were optimised in hanging-drop experiments. Ultimately RPA4515 was crystallised in 2 M ammonium sulphate/0.1 M sodium acetate pH 5.0, after one week incubation at 17 °C. Crystals were flash-cooled in liquid nitrogen in the presence of mother-liquor plus glycerol (20% v/v; added as cryoprotectant), and stored in liquid nitrogen.

Data were collected at the Diamond Light Source (Harwell, UK) on Beamline station I04-1 (2-oxoadipate) and I03 (adipate). Data processing was done with XDS and merging with XSCALE. The crystals of the adipate- and 2-oxoadipate-bound forms were isomorphous. Molecular

replacement was performed with PHASER software, part of the Collaborative Computational Project, Number 4 (CCP4) software suite [55], using Bug27 as a model (PDB accession code 2QPQ). The resultant electron density maps were analysed using COOT [56] and the protein models refined using REFMAC5 [57] software. Validation was performed within COOT and confirmed with MolProbity server [58]. The final model comprises residues 2–300 plus ligands and solvent. The structure factors and co-ordinates can be accessed in the PDB with the accession codes 5OEI (2-oxoadipate bound) and 5OKU (adipate bound). Figures were generated using PYMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC) and Ligplot+ [59].

Acknowledgements

The present work was accomplished with funding from the Brazilian funding agency CNPQ (National Council for Scientific and Technological Development), through a PhD studentship to LTR (248597/2013-2) in the remit of 'science without borders' program.

Conflicts of interest

The authors declare they have no conflicts of interest in connection with the content of this article.

Author contributions

LTR designed and executed most of the experiments, analysed the results and cowrote the paper. JBR and SRD collected protein crystal diffraction data and built the protein structure model. JBR commented on the manuscript. DJK conceived the idea for the project, helped to design the experiments, cowrote and edited the paper and provided supervisory support.

References

1 Reddy VS, Shlykov MA, Castillo R, Sun EI & Saier MH (2012) The major facilitator superfamily (MFS) revisited. *FEBS J* **279**, 2022–2035.
2 Giuliani S, Frank A, Corgliano D, Seifert C, Hauser L & Collart F (2011) Environment sensing and response mediated by ABC transporters. *BMC Genom* **12**, 1–14.
3 Maqbool A, Horler RS, Muller A, Wilkinson AJ, Wilson KS & Thomas GH (2015) The substrate-binding protein in bacterial ABC transporters: dissecting roles in the evolution of substrate specificity. *Biochem Soc Trans* **43**, 1011–1017.
4 Forward JA, Behrendt MC, Wyborn NR, Cross R & Kelly DJ (1997) TRAP transporters: a new family of periplasmic solute transport systems encoded by the *dctPQM* genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *J Bacteriol* **179**, 5482–5493.
5 Winnen B, Hvorup RN & Saier MH Jr (2003) The tripartite tricarboxylate transporter (TTT) family. *Res Microbiol* **154**, 457–465.
6 Hamblin MJ, Shaw JG, Curson JP & Kelly DJ (1990) Mutagenesis, cloning and complementation analysis of C4-dicarboxylate transport genes from *Rhodobacter capsulatus*. *Mol Microbiol* **4**, 1567–1574.
7 Salmon RC, Cliff MJ, Rafferty JB & Kelly DJ (2013) The CouPSTU and TarPQM transporters in *Rhodopseudomonas palustris*: redundant, promiscuous uptake systems for lignin-derived aromatic substrates. *PLoS One* **8**, e59844.
8 Vetting MW, Al-Obaidi N, Zhao S, San Francisco B, Kim J, Wichelecki DJ, Bouvier JT, Solbiati JO, Vu H, Zhang X *et al.* (2015) Experimental strategies for functional annotation and metabolism discovery: targeted screening of solute binding proteins and unbiased panning of metabolomes. *Biochemistry* **54**, 909–931.
9 Fischer M, Zhang QY, Hubbard RE & Thomas GH (2010) Caught in a TRAP: substrate-binding proteins in secondary transport. *Trends Microbiol* **18**, 471–478.
10 Sweet GD, Kay CM & Kay WW (1984) Tricarboxylate-binding proteins of *Salmonella typhimurium*. Purification, crystallization, and physical properties. *J Biol Chem* **259**, 1586–1592.
11 Antoine R, Jacob-Dubuisson F, Drobecq H, Willery E, Lesjean S & Locht C (2003) Overrepresentation of a gene family encoding extracytoplasmic solute receptors in *Bordetella*. *J Bacteriol* **185**, 1470–1474.
12 Ghai R, Rodriguez-Valera F, McMahon KD, Toyama D, Rinke R, Cristina Souza de Oliveira T, Wagner Garcia J, Pellon de Miranda F & Henrique-Silva F (2011) Metagenomics of the water column in the pristine upper course of the Amazon river. *PLoS One* **6**, e23785.
13 Hosaka M, Kamimura N, Toribami S, Mori K, Kasai D, Fukuda M & Masai E (2013) Novel tripartite aromatic acid transporter essential for terephthalate uptake in *Comamonas* sp. strain E6. *Appl Environ Microbiol* **79**, 6148–6155.
14 Piepenbreier H, Fritz G & Gebhard S (2017) Transporters as information processors in bacterial signalling pathways. *Mol Microbiol* **104**, 1–15.
15 Antoine R, Huvent I, Chemlal K, Deray I, Raze D, Locht C & Jacob-Dubuisson F (2005) The periplasmic binding protein of a tripartite tricarboxylate transporter is involved in signal transduction. *J Mol Biol* **351**, 799–809.
16 Herrou J, Bompard C, Antoine R, Leroy A, Rucktooa P, Hot D, Huvent I, Locht C, Villeret V & Jacob-Dubuisson F (2007) Structure-based mechanism of ligand binding for periplasmic solute-binding protein of the Bug family. *J Mol Biol* **373**, 954–964.

17 Huvent I, Belrhali H, Antoine R, Bompard C, Locht C, Jacob-Dubuisson F & Villeret V (2006) Crystal structure of *Bordetella pertussis* BugD solute receptor unveils the basis of ligand binding in a new family of periplasmic binding proteins. *J Mol Biol* **356**, 1014–1026.
18 Huvent I, Belrhali H, Antoine R, Bompard C, Locht C, Jacob-Dubuisson F & Villeret V (2006) Structural analysis of *Bordetella pertussis* BugE solute receptor in a bound conformation. *Acta Crystallogr D Biol Crystallogr* **62**, 1375–1381.
19 Denger K & Cook AM (2010) Racemase activity effected by two dehydrogenases in sulfolactate degradation by *Chromohalobacter salexigens*: purification of (S)-sulfolactate dehydrogenase. *Microbiology* **156**, 967–974.
20 Wubbeler JH, Hiessl S, Schuldes J, Thurmer A, Daniel R & Steinbuchel A (2014) Unravelling the complete genome sequence of *Advenella mimigardefordensis* strain DPN7T and novel insights in the catabolism of the xenobiotic polythioester precursor 3,3′-dithiodipropionate. *Microbiology* **160**, 1401–1416.
21 Larimer FW, Chain P, Hauser L, Lamerdin J, Malfatti S, Do L, Land ML, Pelletier DA, Beatty JT, Lang AS *et al.* (2004) Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. *Nat Biotechnol* **22**, 55–61.
22 Elder DJ & Kelly DJ (1994) The bacterial degradation of benzoic acid and benzenoid compounds under anaerobic conditions: unifying trends and new perspectives. *FEMS Microbiol Rev* **13**, 441–468.
23 Pan C, Oda Y, Lankford PK, Zhang B, Samatova NF, Pelletier DA, Harwood CS & Hettich RL (2008) Characterization of anaerobic catabolism of p-coumarate in *Rhodopseudomonas palustris* by integrating transcriptomics and quantitative proteomics. *Mol Cell Proteomics* **7**, 938–948.
24 McKinlay JB & Harwood CS (2010) Photobiological production of hydrogen gas as a biofuel. *Curr Opin Biotechnol* **21**, 244–251.
25 Heiniger EK, Oda Y, Samanta SK & Harwood CS (2012) How posttranslational modification of nitrogenase is circumvented in *Rhodopseudomonas palustris* strains that produce hydrogen gas constitutively. *Appl Environ Microbiol* **78**, 1023–1032.
26 Xing D, Zuo Y, Cheng S, Regan JM & Logan BE (2008) Electricity generation by *Rhodopseudomonas palustris* DX-1. *Environ Sci Technol* **42**, 4146–4151.
27 Inglesby AE, Beatty DA & Fisher AC (2012) *Rhodopseudomonas palustris* purple bacteria fed *Arthrospira maxima* cyanobacteria: demonstration of application in microbial fuel cells. *RSC Advances* **2**, 4829–4838.
28 Austin S, Kontur WS, Ulbrich A, Oshlag JZ, Zhang W, Higbee A, Zhang Y, Coon JJ, Hodge DB, Donohue TJ *et al.* (2015) Metabolism of multiple aromatic compounds in corn stover hydrolysate by *Rhodopseudomonas palustris*. *Environ Sci Technol* **49**, 8914–8922.
29 Venkidusamy K & Megharaj M (2016) A novel electrophototrophic bacterium *Rhodopseudomonas palustris* strain RP2, exhibits hydrocarbonoclastic potential in anaerobic environments. *Front Microbiol* **7**, 1071.
30 Rinaldi J, Arrar M, Sycz G, Cerutti ML, Berguer PM, Paris G, Estrin DA, Marti MA, Klinke S & Goldbaum FA (2016) Structural insights into the HWE histidine kinase family: the *Brucella* blue light-activated histidine kinase domain. *J Mol Biol* **428**, 1165–1179.
31 Scheepers GH, Lycklama a Nijeholt JA & Poolman B (2016) An updated structural classification of substrate-binding proteins. *FEBS Lett* **590**, 4393–4401.
32 Widenhorn KA, Somers JM & Kay WW (1989) Genetic regulation of the tricarboxylate transport operon (tctI) of *Salmonella typhimurium*. *J Bacteriol* **171**, 4436–4441.
33 Widenhorn KA, Somers JM & Kay WW (1988) Expression of the divergent tricarboxylate transport operon (tctI) of *Salmonella typhimurium*. *J Bacteriol* **170**, 3223–3227.
34 Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL & Liu SQ (2016) Insights into protein-ligand interactions: mechanisms, models, and methods. *Int J Mol Sci* **17**, 144.
35 Fischer M, Hopkins AP, Severi E, Hawkhead J, Bawdon D, Watts AG, Hubbard RE & Thomas GH (2015) Tripartite ATP-independent periplasmic (TRAP) transporters use an arginine-mediated selectivity filter for high affinity substrate binding. *J Biol Chem* **290**, 27113–27123.
36 Harrison FH & Harwood CS (2005) The *pimFABCDE* operon from *Rhodopseudomonas palustris* mediates dicarboxylic acid degradation and participates in anaerobic benzoate degradation. *Microbiology* **151**, 727–736.
37 Donoghue NA & Trudgill PW (1975) The metabolism of cyclohexanol by *Acinetobacte*r NCIB 9871. *Eur J Biochem* **60**, 1–7.
38 Küver J, Xu Y & Gibson J (1995) Metabolism of cyclohexane carboxylic acid by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Arch Microbiol* **164**, 337–345.
39 Harwood CS & Parales RE (1996) The beta-ketoadipate pathway and the biology of self-identity. *Annu Rev Microbiol* **50**, 553–590.
40 Giuliani SE, Frank AM, Corgliano DM, Seifert C, Hauser L & Collart FR (2011) Environment sensing and response mediated by ABC transporters. *BMC Genom* **12** (Suppl 1), S8.
41 Manandhar M & Cronan JE (2017) Pimelic acid, the first precursor of the *Bacillus subtilis* biotin synthesis pathway, exists as the free acid and is assembled by fatty acid synthesis. *Mol Microbiol* **104**, 595–607.

42 Lin S & Cronan JE (2011) Closing in on complete pathways of biotin biosynthesis. *Mol BioSyst* **7**, 1811–1821.
43 van Niel CB (1944) The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria. *Bacteriol Rev* **8**, 1–118.
44 Harwood CS & Gibson J (1988) Anaerobic and aerobic metabolism of diverse aromatic compounds by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Appl Environ Microbiol* **54**, 712–717.
45 Díaz E, Jiménez JI & Nogales J (2013) Aerobic degradation of aromatic compounds. *Curr Opin Biotechnol* **24**, 431–442.
46 Bart JCJ & Cavallaro S (2015) Transiting from adipic acid to bioadipic acid. 1, Petroleum-based processes. *Ind Eng Chem Res* **54**, 1–46.
47 Polen T, Spelberg M & Bott M (2013) Toward biotechnological production of adipic acid and precursors from biorenewables. *J Biotechnol* **167**, 75–84.
48 Kruyer NS & Peralta-Yahya P (2017) Metabolic engineering strategies to bio-adipic acid production. *Curr Opin Biotechnol* **45**, 136–143.
49 Medintz IL & Deschamps JR (2006) Maltose-binding protein: a versatile platform for prototyping biosensing. *Curr Opin Biotechnol* **17**, 17–27.
50 Medintz IL, Goldman ER, Lassman ME & Mauro JM (2003) A fluorescence resonance energy transfer sensor based on maltose binding protein. *Bioconjug Chem* **14**, 909–918.
51 Kim M-K & Harwood CS (1991) Regulation of benzoate-CoA ligase in *Rhodopseudomonas palustris*. *FEMS Microbiol Lett* **83**, 199–203.
52 Beatty JT & Gest H (1981) Biosynthetic and bioenergetic functions of citric acid cycle reactions in *Rhodopseudomonas capsulata*. *J Bacteriol* **148**, 584–593.
53 Guccione E, Del Rocio Leon-Kempis M, Pearson BM, Hitchin E, Mulholland F, Van Diemen PM, Stevens MP & Kelly DJ (2008) Amino acid-dependent growth of *Campylobacter jejuni*: key roles for aspartase (AspA) under microaerobic and oxygen-limited conditions and identification of AspB (Cj0762), essential for growth on glutamate. *Mol Microbiol* **69**, 77–93.
54 Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H *et al.* (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci USA* **103**, 15835–15840.
55 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **67**, 235–242.
56 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
57 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53**, 240–255.
58 Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D: Biol Crystallogr* **66**, 12–21.
59 Laskowski RA & Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778–2786.