STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Structure of an ABC transporter solute-binding protein specific for the amino sugars glucosamine and galactosamine

Umesh Yadava,[a] Matthew W. Vetting,[b] Nawar Al Obaidi,[b] Michael S. Carter,[c] John A. Gerlt[c] and Steven C. Almo[b]*

[a]Department of Physics, DDU Gorakhpur University, Gorakhpur 273 009, India, [b]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY 10461, USA, and [c]Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA. *Correspondence e-mail: steve.almo@einstein.yu.edu

Received 20 January 2016
Accepted 4 May 2016

Edited by M. L. Pusey, University of Alabama, USA

Keywords: solute-binding protein; differential scanning fluorimetry; Thermofluor; *Agrobacterium vitis*; Avi_5305; Pfam13407.

PDB references: Avi_5305, complex with D-galactosamine, 5br1; complex with D-glucosamine, 4y9t

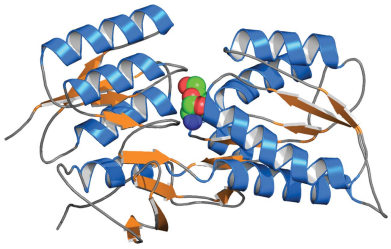
Supporting information: this article has supporting information at journals.iucr.org/f

The uptake of exogenous solutes by prokaryotes is mediated by transport systems embedded in the plasma membrane. In many cases, a solute-binding protein (SBP) is utilized to bind ligands with high affinity and deliver them to the membrane-bound components responsible for translocation into the cytoplasm. In the present study, Avi_5305, an *Agrobacterium vitis* SBP belonging to Pfam13407, was screened by differential scanning fluorimetry (DSF) and found to be stabilized by D-glucosamine and D-galactosamine. Avi_5305 is the first protein from Pfam13407 shown to be specific for amino sugars, and co-crystallization resulted in structures of Avi_5305 bound to D-glucosamine and D-galactosamine. Typical of Pfam13407, Avi_5305 consists of two α/β domains linked through a hinge region, with the ligand-binding site located in a cleft between the two domains. Comparisons with *Escherichia coli* ribose-binding protein suggest that a cation–π interaction with Tyr168 provides the specificity for D-glucosamine/D-galactosamine over D-glucose/D-galactose.

1. Introduction

The great success of genome-sequencing projects has resulted in an enormous number of protein sequences for which functional annotations have been assigned on a computational basis (*i.e.* sequence similarity), resulting in a large proportion of sequences that are either mis-annotated or under-annotated (Gerlt *et al.*, 2011). To address this challenge, new strategies are needed to elucidate individual enzyme activities and to define *in vivo* functions and new metabolic pathways. The first reactant in a metabolic pathway is often a solute in the extracellular milieu that is moved into the cytoplasm by a high-affinity transport system. The identification of this first reactant can provide a significant insight into the function of co-regulated and/or co-localized genes. In many transport systems, a soluble periplasmic or membrane-bound solute-binding protein (SBP) is utilized to bind and deliver their cognate ligands to membrane-bound translocator subunits. Recent work has demonstrated that differential scanning fluorimetry screening against focused metabolite libraries is an effective approach for identifying SBP ligands, resulting in functional annotation of transporters and their associated metabolic pathways (Vetting *et al.*, 2015; Huang *et al.*, 2015; Wichelecki *et al.*, 2015).

ATP-binding cassette (ABC) transporters, a large class of solute transporters, are typically composed of two nucleotide-binding domains, two transmembrane domains and an SBP, with solute transport being coupled to ATP hydrolysis (Locher, 2009; Berntsson *et al.*, 2010; Oldham *et al.*, 2008).



Avi_5305, an example of an ABC transporter-associated SBP from the plant pathogen *Agrobacterium vitis* S4, belongs to Pfam family Pfam13407, the members of which typically bind monosaccharides such as ribose, galactose and arabinose (http://pfam.xfam.org/family/PF13407; Finn *et al.*, 2016). As the causative agent of crown gall in the trunks of grapevines, *A. vitis* reduces crop yields by decreasing vine strength and growth (Burr & Otten, 1999; Schroth, 1988). Given the evidence that carbohydrate accumulation and degradation are associated with pathogenicity and the production of crown-gall tumors (Conner *et al.*, 1937; Aloni & Ullrich, 2008; Shimoda *et al.*, 1990; Hodgson *et al.*, 1945), studies of sugar transport in *A. vitis* might provide new agricultural opportunities.

The present study provides differential scanning fluorimetry results and three-dimensional structural descriptions of the specific binding of D-glucosamine and D-galactosamine by Avi_5305. D-Glucosamine is the monomeric unit of chitin (Carlstrom, 1957) and is one of the monomers in heparin (Gatti *et al.*, 1979). The N-acetylated version is a component of Gram-negative outer membrane lipopolysaccharides in bacterial cell walls, and is a common constituent saccharide unit in the glycoproteins presented on the surfaces of cells from multiple domains of life. Galactosamine is secreted by some fungi (Distler & Roseman, 1960), and its N-acetylated version is also commonly found in glycoproteins (Lis & Sharon, 1993). As nitrogen and/or carbon sources in bacteria, glucosamine and galactosamine are commonly phosphorylated, deaminated and isomerized to a hexose-6-phosphate, which can be cleaved to yield glycolytic intermediates (Wolfe *et al.*, 1957; Nakada & Wolfe, 1956; Morita *et al.*, 1956; Brinkkotter *et al.*, 2000). There are currently no data available indicating that grapevines secrete amino sugars and that their utilization would be advantageous to *A. vitis*; however, glucosamine has been shown to inhibit crown galls caused by the related bacterium *A. tumefaciens* (Richardson & Morre, 1978).

2. Methods

2.1. Cloning, expression and purification of Avi_5305

A PCR amplicon of the *avi_5305* gene (residues 24–346, excluding the N-terminal periplasmic signal sequence) was produced using *A. vitis* S4 genomic DNA with the forward primer 5′-TACTTCCAATCCATGGCCCAAACAAAAGGGATGGTTTATTAC-3′ and the reverse primer 5′-CGACGCTAGTTTAGGCAGCTCATTGTCATTTCCACCTAT-3′ (ligation-independent cloning sites are underlined). The PCR amplicon was cloned into the N-terminal TEV-cleavable 6×His-tag vector pNIC23-Bsa4, a pET-23-based variant of the pNIC28-Bsa4 vector (Savitsky *et al.*, 2010), by ligation-independent cloning (LIC; Aslanidis & de Jong, 1990). All growth media contained 100 µg ml^{-1} carbenicillin and 34 µg ml^{-1} chloramphenicol. The vector containing the cloned target was transformed into *Escherichia coli* BL21 (DE3) cells containing the pRIL plasmid (Stratagene) and used to inoculate a 20 ml culture of 2×YT medium. After overnight growth, the culture was used to inoculate 2 l selenomethionine-containing (or methionine-containing for the native) ZYP-5052 autoinduction medium (Studier, 2005) in a LEX 48 airlift fermenter, which was incubated for 4 h at 37°C and then for an additional 12–16 h at 25°C. The cells were pelleted by centrifugation and stored at −80°C. The cells were lysed by sonication in lysis buffer (1 g of cells per 3 ml of 50 m*M* Tris pH 8.0, 150 m*M* NaCl, 20 m*M* imidazole, 0.5% Tween 20). The lysate was clarified by centrifugation and loaded onto a His60 Ni Resin column (Clontech) pre-equilibrated with buffer *A* (50 m*M* Tris pH 8.0, 150 m*M* NaCl, 20 m*M* imidazole). The bound proteins were washed with ten column volumes of buffer *A* and eluted with buffer *B* (50 m*M* Tris pH 8.0, 150 m*M* NaCl, 400 m*M* imidazole). The protein was further purified using a HiLoad Superdex 200 16/60 pg gel-filtration column (GE Healthcare) equilibrated with buffer *C* (20 m*M* Tris pH 8.0, 150 m*M* NaCl, 5 m*M* DTT). Proteins were analyzed for purity by SDS–PAGE and the concentration was determined from the absorbance at 280 nm (Abs 0.1% = 0.880). Engineered TEV protease (Blommel & Fox, 2007) was added to purified proteins in a 1:80 ratio and the sample was incubated on ice for 2 h and then buffer-exchanged into 20 m*M* Tris pH 8.0, 5 m*M* DTT by dilution and ultrafiltration centrifugation. The sample was concentrated to 40–60 mg ml^{-1}, flash-cooled in liquid N_2 and stored at −80°C.

2.2. Differential scanning fluorimetry of Avi_5305

The purified protein was screened by differential scanning fluorometry (DSF) in 384-microwell plates using an Applied Biosystems 7900 HT real-time PCR system with excitation at 490 nm and emission at 530 nm. The final assay mixture (10 µl) consisted of 10 µ*M* protein, 1 m*M* ligand and 5× SYPRO Orange (Thermo Fisher) in 100 m*M* HEPES pH 7.5, 150 m*M* NaCl, 5 m*M* DTT. The screening library consisted of 405 metabolites (Wichelecki *et al.*, 2015), some as mixtures of up to six compounds, with each condition in duplicate. The temperature was increased from 22 to 99°C at 3°C min^{-1}, with the melting temperature of the protein (T_m) calculated by fitting the melting curve to the Boltzmann equation (Niesen *et al.*, 2007). ΔT_m values were calculated as the difference of the average T_m values measured with ligand and without ligand (control wells).

2.3. Crystallization, data collection and structure determination of Avi_5305

Avi_5305 was crystallized by sitting-drop vapor diffusion in 96-well Intelli-Plates (Art Robbins) stored at 18°C; the initial crystal hits originated from the commercial screens MCSG1–4 (Microlytic). The final crystallization conditions consisted of 0.5 µl protein solution (40–60 mg ml^{-1}, 10 m*M* D-galactosamine or D-glucosamine) combined with 0.5 µl reservoir solution (30–40% PEG 4000). Crystals were mounted on nylon loops and flash-cooled by plunging them directly into liquid nitrogen without added cryoprotectant. The structure of the Avi_5305–D-glucosamine complex (PDB entry 4y9t) was determined using a data set collected from a selenomethionine-

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution bin.

PDB entry	4y9t	5br1
Co-crystallized ligand	D-Glucosamine	D-Galactosamine
Data-collection statistics		
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 36.63, b = 63.61, c = 127.39	a = 36.78, b = 63.70, c = 119.58
Resolution (Å)	100–1.80 (1.83–1.80)	16.6–1.85 (1.88–1.85)
R_{merge}	0.135 (0.740)	0.127 (0.608)
No. of unique reflections	25086 (1237)	21553 (782)
Multiplicity	13.3 (13.6)	4.9 (2.8)
Completeness (%)	88.2 (87.1)	87.0 (60.8)
Mean $I/\sigma(I)$	19.8 (3.75)	10.1 (2.20)
Wilson B factor (Å^2)	18.5	21.9
Refinement statistics		
Resolution (Å)	24.0–1.80 (1.87–1.80)	16.5–1.85 (1.95–1.85)
R_{work} (%)	14.8 (19.0)	18.3 (28.8)
R_{free} (5% of data) (%)	18.9 (25.9)	28.3 (32.2)
No. of residues	321	320
No. of waters	319	238
Total No. of atoms	3049	2981
B factor (Å^2)		
Protein	23.1	31.6
Waters	29.1	31.3
Ligand	11.2	24.1
R.m.s.d., bond lengths (Å)	1.32	1.40
R.m.s.d., angles (°)	0.011	0.013
Ramachandran outliers (%)	0.3	0.3
Side-chain outliers (%)	1.2	0.4

containing crystal on beamline 31-ID (LRL-CAT) at the Advanced Photon Source using a wavelength of 0.9793 Å at 100 K, while the structure of the Avi-5305–D-galactosamine complex (PDB entry 5br1) was obtained from data collected from a methionine-containing crystal at 100 K with an R-AXIS IV^{++} detector using Cu $K\alpha$ radiation generated by an RU-300 generator. Data were integrated and scaled using *HKL*-3000 (Minor *et al.*, 2006). Initial phases were determined by selenomethionine SAD with *SHELX* in *HKL*-3000 (Sheldrick, 2008) and an initial model was built using *ARP*/*wARP* in *HKL*-3000 (Morris *et al.*, 2003). Iterative cycles of manual rebuilding with the molecular-graphics program *Coot* (Emsley *et al.*, 2010) and refinement against the data with *PHENIX* (Adams *et al.*, 2010) were performed until convergence was achieved. During the final refinement cycles, ligands were built into the observed difference density and TLS refinement (Winn *et al.*, 2001) was performed with TLS ranges as determined within *PHENIX*. Data-collection and refinement statistics are given in Table 1. Diffraction from the Avi-5305–D-galactosamine complex crystal exhibited streaky intensities, which is the most likely cause of the large difference in R and R_{free} for the refined structure.

3. Results

The N-terminal 23 amino acids, which were predicted to be a periplasmic secretion signal by *SignalP* (Bendtsen *et al.*, 2004), were not included in cloning. Avi_5305 was purified to homogeneity by metal-affinity chromatography followed by size-exclusion chromatography, with migration on SDS–PAGE consistent with the expected mass of 36.8 kDa (Fig. 1). DSF screening of Avi_5305 with a 405-component ligand library (for details, see Supplemental Table 1 of Wichelecki *et al.*, 2015) consisting of amino acids, sugars, benzoates and other environmentally available compounds demonstrated significant thermal stabilization of Avi_5305 by a D-glucosamine/*N*-acetyl-D-glucosamine cocktail and a D-galactosamine/*N*-acetyl-D-galactosamine cocktail, with ΔT_m values of 11.5 and 7.5°C, respectively. In contrast, a D-glucose/L-glucose cocktail and a D-galactose/L-galactose cocktail yielded ΔT_m values of 0.2 and −0.2°C, respectively, suggesting the importance of the amine functionality. Deconvolution of the amino-sugar cocktails yielded stabilization by D-glucosamine and D-galactosamine (13.6 *versus* 9.9°C; Fig. 2), while Avi_5305 was not stabilized by the acetylated derivatives (ΔT_m of 0.7 *versus* −0.4°C).

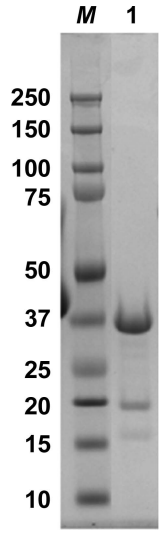


Figure 1
SDS–PAGE analysis of purified Avi_5305. Shown are Avi_5305 after purification (lane 1) and molecular standards with molecular weights labeled in kDa (lane *M*).

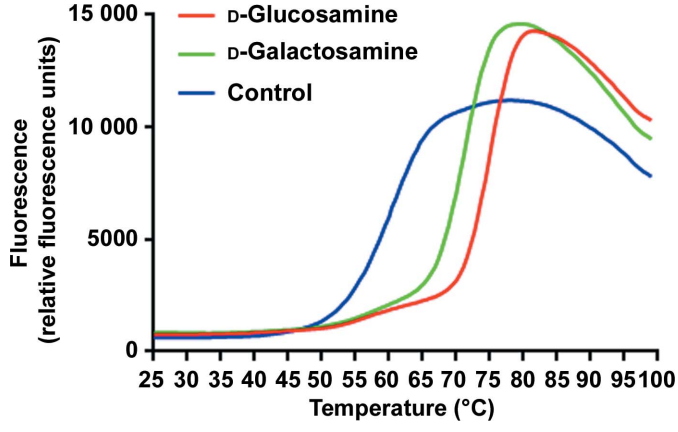


Figure 2
Differential scanning fluorometry (DSF) of Avi_5305. DSF curves are shown for Avi_5305 with no added ligand (blue) and with either 1 m*M* D-glucosamine (red) or D-galactosamine (green).

Based on the results of DSF screening, co-crystallization of the protein was performed with *N*-acetyl-D-glucosamine, *N*-acetyl-D-galactosamine, D-glucosamine and D-galactosamine, with only the D-glucosamine and D-galactosamine complexes yielding well diffracting crystals. Both complexes crystallized in the orthorhombic space group $P2_12_12_1$, with one molecule in the asymmetric unit. The entire sequence was fitted to electron density, except for residues 24–26, and clear difference electron density was observed for the co-crystallized ligands (Fig. 3*a*). Details of data-collection, structure-determination and refinement statistics are listed in Table 1.

The structure of Avi_5305 consists of two α/β domains, each of which comprises a central β-sheet sandwiched between α-helices (Fig. 3*b*). Domain A is composed of residues 27–130 and 270–317, while domain B is composed of residues 132–268, with the cross-over segments between the domains acting as a hinge. Additionally, the C-terminal residues 318–346 interact with both domains and provide support to the hinge region. Berntsson *et al.* (2010) classified SBPs into six clusters based on structural similarity, where Avi_5305 would be classified as a cluster B SBP with its characteristic three cross-over hinges connecting the two domains. Cluster B consists of class I SBPs (as defined by Fukami-Kobayashi *et al.*, 1999) that bind carbohydrates, branched-chain amino acids, natriuretic peptides and autoinducer 2. Cluster B also includes the effector-binding domains of *lac*-repressor type transcription factors, for example the purine repressor PurR (Schumacher *et al.*, 1994). The most structurally similar protein, as calculated by the *PDBeFold* server (Krissinel & Henrick, 2004; http://www.ebi.ac.uk/msd-srv/ssm), is the *E. coli* ribose-binding protein (*Ec*RBP; PDB entry 1drk; r.m.s.d. 1.40 Å) bound to ribose (in the pyranose form), along with other pyranose-binding proteins from Pfam13047. Previously determined Pfam13407 structures exhibit low sequence identity to Avi_5305 (<20%); therefore, the structure of Avi_5305 will enhance efforts to map ligand specificities within the solute-binding family. Examination of molecular packing suggests that Avi_5305 is monomeric, which is typical of the great majority of solute-binding proteins.

Solute-binding proteins from the ABC transporter family utilize a 'Venus flytrap' mechanism (Mao *et al.*, 1982), with a large conformational change between the two α/β domains upon interaction with the cognate ligand leading to encapsulation of the ligands. The structure determinations of Avi_5305

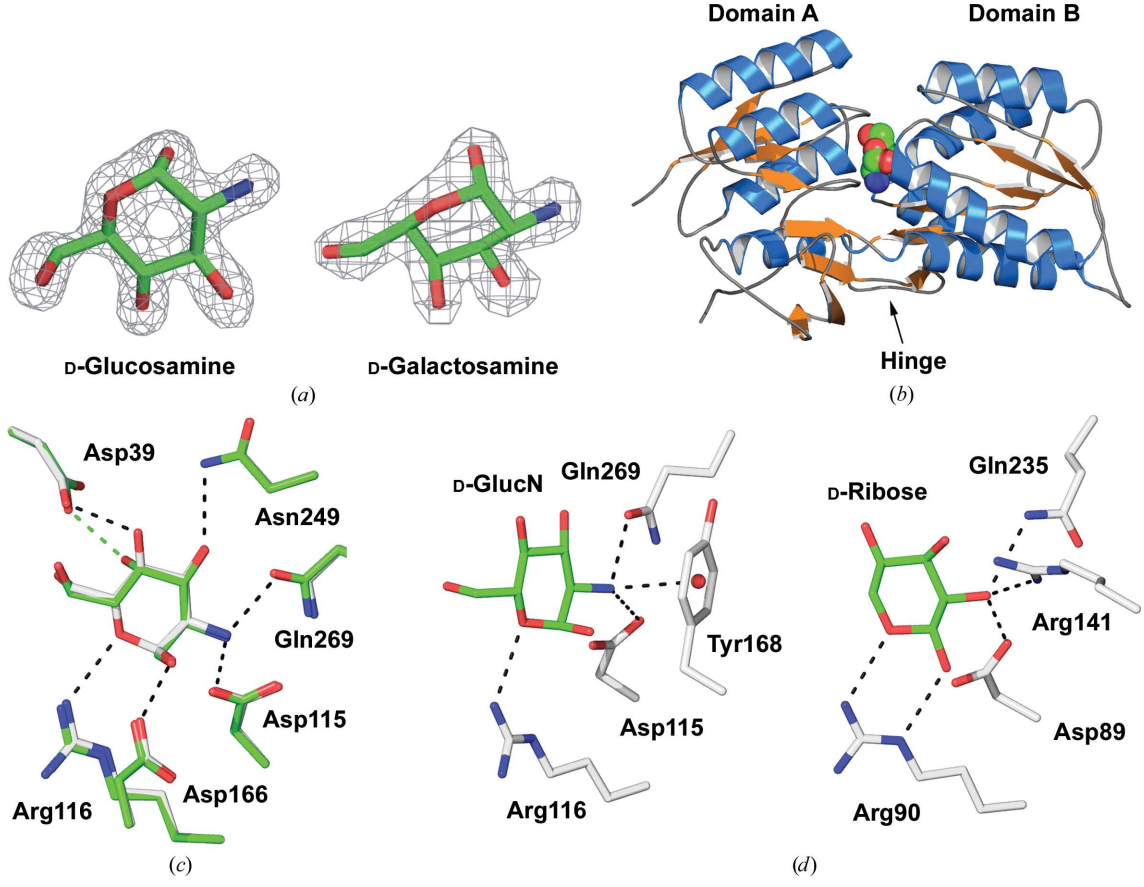


Figure 3
Crystal structures of Avi_5305 bound to D-glucosamine and D-galactosamine. (*a*) Difference OMIT maps of the D-glucosamine and D-galactosamine complexes contoured at 2.0σ prior to inclusion of the ligand. (*b*) Ribbon diagram of Avi_5305 in complex with D-glucosamine. (*c*) Superposition of the Avi_5305 complexes with D-glucosamine (white C atoms; PDB entry 4y9t) and D-galactosamine (green C atoms; PDB entry 5br1). The interaction of Avi_5305 with Tyr168 is not included for clarity; see (*d*). (*d*) A comparison of Avi_5305 in complex with D-glucosamine (left) and *E. coli* ribose-binding protein in complex with ribose (PDB entry 1drk; right). Shown are a conserved arginine interaction with the endocyclic O atom and the similarities and differences in the coordination of the C2 functional group. A sphere, 3.1 Å from the amine group, marks the centroid of the phenoxy ring of Tyr168.

demonstrate the binding of α-D-glucosamine and α-D-galactosamine within a buried cleft composed of residues from both domains. The structures of Avi_5305 with D-glucosamine and D-galactosamine are highly similar (r.m.s.d. of 0.29 Å over 316 C^{α} atoms), with conserved hydrogen-bonding interactions with the side chains of Arg116, Asp166, Gln269, Asn249 and Asp39 (Fig. 3*c*). Despite the difference in stereochemistry at C4, these hydroxyls in D-glucosamine and D-galactosamine are coordinated by Asp39 in both complexes. Similar to those observed for the binding of ribose by *Ec*RBP, the pyranose ring forms stacking interactions with spatially conserved aromatics: Phe41 and Trp196 in the case of Avi_5305 and Phe15 and Phe164 in the case of *Ec*RBP (not shown). Arg116, a conserved arginine in *Ec*RBP and Avi_5305, makes a hydrogen bond to the endocyclic O atom in both structures and therefore may be vital for the recognition of pyranose ligands. The pyranose amine of Avi_5305 is coordinated by the side chain of Asp115 and Gln269 OE1 and forms a cation–π interaction with Tyr168 (Fig. 3*d*). Interestingly, in *Ec*RBP the corresponding Asp and Gln NE2 are conserved; however, the spatial equivalent of Tyr168 is replaced by an arginine (Arg141), suggesting that the cation–π interaction with Tyr168 is likely to be a major determinant for the binding of D-glucosamine/D-galactosamine relative to ligands with hydroxyl groups such as D-glucose/D-galactose.

4. Discussion

The amino sugars glucosamine and galactosamine are abundant in soil organic matter (Zhang *et al.*, 2013) and marine environments (Benner & Kaiser, 2003). Chitin, the second most abundant polysaccharide after cellulose, is a structural polymer of *N*-acetylglucosamine produced by a number of organisms including crustaceans, insects and fungi (Carlstrom, 1957; Latgé, 2007; Muzzarelli *et al.*, 1986). Amino sugars are important structural components of prokaryotic cell walls, where they occur in peptidoglycan, lipopolysaccharides and pseudopeptidoglycan (Brock *et al.*, 1994). The N-acetylated versions of amino sugars are also commonly found in eukaryotic glycoproteins (Lis & Sharon, 1993).

In prokaryotes, for the few systems studied, glucosamine transport occurs *via* sugar-transporting phosphotransferase systems (PTSs), which do not utilize SBPs (Gaugué *et al.*, 2013; Uhde *et al.*, 2013; Curtis & Epstein, 1975). Here, differential scanning fluorimetry and crystallography demonstrated that an *A. vitis* S4 SBP (Avi_5305) from Pfam13407 was specific for glucosamine and galactosamine. Recently, Zhao & Binns (2014) described a Pfam13407 SBP-containing ATP transporter (GxySBA) from *A. tumefaciens* C58 (Atu3576) that was involved in the utilization of a wide range of sugars including glucose, xylose and glucosamine. While *A. vitis* S4 contains an ortholog of Atu3576 (Avi_1212, 77% sequence identity), *A. tumefaciens* C58 does not contain an ortholog of the amino sugar-specific SBP Avi_5305 (sequence identity <35%). Similar to the *A. tumefaciens* C58 transporter GxySBA, the genomic environment of Avi_5305 contains genes for an ATP-binding subunit (Avi_5304), a transmembrane subunit (Avi_5307) and a ROK-family transcriptional regulator (Avi_5303) with a helix–turn–helix (HTH) DNA-binding domain.

Avi_5305 has few potential orthologs based on sequence (five sequences with sequence identity of >40%), which makes genome-context comparisons difficult, although the residues making contact with the ligand are conserved amongst them. Putative orthologous gene clusters in four *Mesorhizobium* species encode a short-chain dehydrogenase (SDH) (Pfam00106), which is also clustered with *avi_5305* in *A. vitis* (Avi_5308) and could be involved in oxidation/reduction of the ligand after transport. In addition, encoded proximal to the SDH gene in *M. ciceri* is a GCN5 *N*-acetyltransferase (Mesci_4407) that could produce acetylated derivatives from the transported ligand. The assimilation of glucosamine and galactosamine is most often not catalyzed by identical enzymes, despite the similarity in the strategies for their degradation. Typically, they are phosphorylated, isomerized and deaminated to make a hexose phosphate that can be cleaved into three-carbon glycolytic intermediates (Nakada & Wolfe, 1956; Wolfe *et al.*, 1957; Morita *et al.*, 1956; Brinkkotter *et al.*, 2000). Assimilation of the glucosamine that is imported *via* Avi_5305 is likely to be catalyzed by Avi_6235 (which is annotated as a glucosamine:fructose-6-phosphate aminotransferase), given that *avi_6235* is clustered with genes annotated to encode enzymes for the conversion of *N*-acetylglucosamine to fructose-6-phosphate, a glycolytic intermediate. A galactosamine-degradation pathway, however, is less evident, as homologs of enzymes involved in characterized galactosamine-degradation pathways (Brinkkotter *et al.*, 2000) are not present in *A. vitis*. While the specificity of Avi_5305 suggests that the associated transporter has a major role in amino-sugar transport in *A. vitis* S4, physiological studies are required to verify its function *in vivo*.

Differential scanning fluorimetry (DSF) has emerged as a facile technique for the high-throughput screening of new ligands for SBPs, resulting in the discovery of new catabolic pathways for ethanolamine, D-threitol, L-threitol, erythritol, D-altritol and galactitol (Vetting *et al.*, 2015; Giuliani *et al.*, 2008; Michalska *et al.*, 2012; Huang *et al.*, 2015; Wichelecki *et al.*, 2015). Using DSF and a metabolite ligand library, an SBP from *A. vitis* S4 (Avi_5305) was found to be specific for the amino sugars D-glucosamine and D-galactosamine. Avi_5305 is the first protein from Pfam13407 which has been found to be specific for amino sugars. Crystal structures of Avi_5305 with D-glucosamine and D-galactosamine highlight similarities to and differences from the most structurally similar protein, *E. coli* ribose-binding protein, and highlight a cation–π interaction with the amine functionality as a binding determinant.

Acknowledgements

UY thanks the University Grants Commission, New Delhi, India for the award of a Raman Fellowship for postdoctoral studies in the USA. This work was supported by the US National Institutes of Health (U54GM093342). Use of the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE) Office of

Science by Argonne National Laboratory, was supported by the US DOE under Contract No. DE-AC02-06CH11357. Use of the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source was provided by Eli Lilly Company, which operates the facility.

References

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

Aloni, R. & Ullrich, C. I. (2008). *Agrobacterium: From Biology to Biotechnology*, edited by T. Tzfira & V. Citovsky, pp. 565–591. New York: Springer.

Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.

Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004). *J. Mol. Biol.* **340**, 783–795.

Benner, R. & Kaiser, K. (2003). *Limnol. Oceanogr.* **48**, 118–128.

Berntsson, R. P., Smits, S. H., Schmitt, L., Slotboom, D. J. & Poolman, B. (2010). *FEBS Lett.* **584**, 2606–2617.

Blommel, P. G. & Fox, B. G. (2007). *Protein Expr. Purif.* **55**, 53–68.

Brinkkotter, A., Kloss, H., Alpert, C. & Lengeler, J. W. (2000). *Mol. Microbiol.* **37**, 125–135.

Brock, T. D., Madigan, M. T., Martinko, J. M. & Parker, J. (1994). *Biology of Microorganisms*, 7th ed. Englewood Cliffs: Prentice Hall.

Burr, T. J. & Otten, L. (1999). *Annu. Rev. Phytopathol.* **37**, 53–80.

Carlstrom, D. (1957). *J. Biophys. Biochem. Cytol.* **3**, 669–683.

Conner, H. A., Riker, A. J. & Peterson, W. H. (1937). *J. Bacteriol.* **34**, 221–236.

Curtis, S. J. & Epstein, W. (1975). *J. Bacteriol.* **122**, 1189–1199.

Distler, J. J. & Roseman, S. (1960). *J. Biol. Chem.* **235**, 2538–2541.

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

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J. & Bateman, A. (2016). *Nucleic Acids Res.* **44**, D279–D285.

Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. (1999). *J. Mol. Biol.* **286**, 279–290.

Gatti, G., Casu, B., Hamer, G. K. & Perlin, A. S. (1979). *Macromolecules*, **12**, 1001–1007.

Gaugué, I., Oberto, J., Putzer, H. & Plumbridge, J. (2013). *PLoS One*, **8**, e63025.

Gerlt, J. A., Allen, K. N., Almo, S. C., Armstrong, R. N., Babbitt, P. C., Cronan, J. E., Dunaway-Mariano, D., Imker, H. J., Jacobson, M. P., Minor, W., Poulter, C. D., Raushel, F. M., Sali, A., Shoichet, B. K. & Sweedler, J. V. (2011). *Biochemistry*, **50**, 9950–9962.

Giuliani, S. E., Frank, A. M. & Collart, F. R. (2008). *Biochemistry*, **47**, 13974–13984.

Hodgson, R., Riker, A. & Peterson, W. (1945). *J. Biol. Chem.* **158**, 89–100.

Huang, H., Carter, M. S., Vetting, M. W., Al-Obaidi, N., Patskovsky, Y., Almo, S. C. & Gerlt, J. A. (2015). *J. Am. Chem. Soc.* **137**, 14570–14573.

Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.

Latgé, J.-P. (2007). *Mol. Microbiol.* **66**, 279–290.

Lis, H. & Sharon, N. (1993). *Eur. J. Biochem.* **218**, 1–27.

Locher, K. P. (2009). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 239–245.

Mao, B., Pear, M. R., McCammon, J. A. & Quiocho, F. A. (1982). *J. Biol. Chem.* **257**, 1131–1133.

Michalska, K., Chang, C., Mack, J. C., Zerbs, S., Joachimiak, A. & Collart, F. R. (2012). *J. Mol. Biol.* **423**, 555–575.

Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* D**62**, 859–866.

Morita, R. Y., Nakada, H. I. & Wolfe, J. B. (1956). *Arch. Biochem. Biophys.* **64**, 480–488.

Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.

Muzzarelli, R. A. A., Jeuniaux, C. & Gooday, G. W. (1986). *Chitin in Nature and Technology*. New York: Plenum Press.

Nakada, H. I. & Wolfe, J. B. (1956). *Arch. Biochem. Biophys.* **64**, 489–497.

Niesen, F. H., Berglund, H. & Vedadi, M. (2007). *Nature Protoc.* **2**, 2212–2221.

Oldham, M. L., Davidson, A. L. & Chen, J. (2008). *Curr. Opin. Struct. Biol.* **18**, 726–733.

Richardson, C. L. & Morre, D. J. (1978). *Int. J. Plant. Sci.* **139**, 196–201.

Savitsky, P., Bray, J., Cooper, C. D., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A. & Gileadi, O. (2010). *J. Struct. Biol.* **172**, 3–13.

Schroth, M. N. (1988). *Plant Dis.* **72**, 241–246.

Schumacher, M. A., Choi, K. Y., Zalkin, H. & Brennan, R. G. (1994). *Science*, **266**, 763–770.

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

Shimoda, N., Toyoda-Yamamoto, A., Nagamine, J., Usami, S., Katayama, M., Sakagami, Y. & Machida, Y. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 6684–6688.

Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.

Uhde, A., Youn, J.-W., Maeda, T., Clermont, L., Matano, C., Krämer, R., Wendisch, V. F., Seibold, G. M. & Marin, K. (2013). *Appl. Microbiol. Biotechnol.* **97**, 1679–1687.

Vetting, M. W. *et al.* (2015). *Biochemistry*, **54**, 909–931.

Wichelecki, D. J., Vetting, M. W., Chou, L., Al-Obaidi, N., Bouvier, J. T., Almo, S. C. & Gerlt, J. A. (2015). *J. Biol. Chem.* **290**, 28963–28976.

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

Wolfe, J. B., Britton, B. B. & Nakada, H. I. (1957). *Arch. Biochem. Biophys.* **66**, 333–339.

Zhang, B., Yang, X., Drury, C. F. & Zhang, X. (2013). *Soil Sci. Soc. Am. J.* **77**, 842–849.

Zhao, J. & Binns, A. N. (2014). *J. Bacteriol.* **196**, 3150–3159.