PURIFICATION AND SECONDARY STRUCTURE ANALYSIS OF MAJOR FACILITATOR SUPERFAMILY TRANSPORT PROTEIN GAB15256 FROM *ARTHROBACTER GLOBIFORMIS*

Irshad Ahmad[1,2], Sher Bahadar Khan[3], Nighat Nawaz[4], Mohammad Zahid Mustafa[5]*, Asadullah[5], Mutiullah Khattak[2], Momin Khan[2], Asif Ali[2], Yasar Mehmood Yousafzai[2], Sadeeq ur Rehman[6] and Chris Thompson[7]

[1]*School of biomedical Sciences, University of Leeds, United Kingdom*
[2]*Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan*
[3]*Department of Animal Health, The University of Agriculture, Peshawar, Pakistan*
[4]*Department of Chemistry, Islamia College Peshawar, Pakistan*
[5]*Center for Advanced Studies in Vaccinology & Biotechnology (CASVAB), University of Balochistan, Quetta, Pakistan*
[6]*College of Animal Husbandry and Veterinary Sciences, Abdul Wali Khan University, Mardan, Pakistan*
[7]*School of Biosciences, University of Birmingham, United Kingdom*
*** Corresponding author:** Phone: +92-3342439980; Email: zunzah@gmail.com

ABSTRACT

Transport protein GAB15256 from *Arthrobacter globiformis* belongs to the Major Facilitator Superfamily (MFS) family. *In silico* analysis of this protein is myo-inositol transporter but specific biological function is unknown yet. Therefore, large scale production and purification would help to study the exact role of GAB15256 protein. In this study, GAB15256 gene was cloned and over expressed in *E. coli* BL21 (DE3) strain. Membranes expressing GAB15256 were solubilised with 1% *n*-dodecyl-*β*-D-maltoside (DDM) and were purified by immobilised metal affinity chromatography using a Ni-NTA resin. Far UV circular dichroism spectroscopy demonstrates that purified protein GAB15256 was largely alpha helical secondary structure and has retained normal folding after passing through the various steps of purification. The thermal stability of GAB15256 was analysed by ramping the temperature from 5-90 ºC and finally back to 5 ºC. On returning the temperature to 5 ºC, there was no evidence that suggests refolding of the protein and a melting temperature of 45.8 ºC was estimated using Global Analysis CD software 3.

Key words: Bacterial protein, Membrane protein overexpression, Purification, CD analysis

INTRODUCTION

Arthrobacters are widely distributed in the environment, typically found in soil. *Arthrobacter globiformis* is a Gram-negative bacterium species from the genus *Arthrobacter* (Jones and Keddie, 1992). A number of bacteria possesses in their genome annotated ioL genes for use of inositol but very limited studies are available on myo-inositol metabolism and its importance for the cell. Many microorganisms, including *Arthrobacter globiformis* are able to use the naturally occurring myo-inositol as their sole carbon and energy source (Yoshida *et al*., 1997; Berman and Magasanik, 1966; Kawsar *et al*., 2004; Krings *et al*., 2006; Poole *et al*., 1994). *Corynebacterium glutamicum* also utilises myo-inositol as a carbon and energy source (Krings *et al*., 2006; Tatusova *et al*., 2016). Generally in chemically complex settings, bacterial growth involves multiple systems for the passage of metabolic substances. In the kingdom of bacteria, the uptake of various molecular substrates as anabolic building blocks and energy sources are facilitated by specific transmembrane transporters (Boutte *et al.*, 2008). The genus *Arthrobacter* transports and metabolises a variety of biological nutrients, comprising the copious ecological carbohydrate myo-inositol (Boutte *et al*., 2008). The GAB15256 protein from *Arthrobacter globiformis* is transporter of myo-inositol and is classified as a member of the MFS family. Topology of this protein is predicted to compromise 12 transmembrane helices. The wild type GAB15256 protein consist of 480 amino acids with the C and N termini situated in the cytoplasm, which allow addition of the hexahistidine tag. A carbocyclic sugar myo-inositol, or simply inositol is abundant in different tissues of mammalian especially brain and in response to a variety of hormones, mediates cell signal transduction neurotransmitters, osmoregulation and growth factors (Parthasarathy *et al*., 2006). Also, in many microbial ecosystems, inositol is very significant molecule comprising of sugar alcohol with half the sweetness of sucrose. In bacteria and archaea inositol has role as antioxidants, as cell membrane components, as carbon storage units and as osmolytes (Galbraith *et al.*, 1998; Roberts, 2006). Inositol can serve as source of phosphorus in the phosphorylated forms (Mullaney *et al*., 2000; Turner *et al*., 2002). The function of inositol

phosphates in bacteria is not well understood, but in eukaryotes, this plays an important role in regulation of membrane traffic and signal transduction at the cell surface the cytoskeleton, membranes permeability and transport functions. *In silico* analysis of this protein is myoinositol transporter but specific biological function is unknown yet. Therefore, large scale production and purification would help to study the exact role of GAB15256 protein. In this study we establish procedures for the recombinant protein expression of GAB15256 gene that was cloned and over expressed in *E. coli* BL21 (DE3) (Robinson, 2011). Purification of GAB15256 protein enables initial studies of structure and function *in vitro*. The results of this study shall hopefully lay the foundation for the structural and functional characterisation of uncharacterized membrane protein of MFS family.

MATERIALS AND METHODS

Cloning of GAB15256 gene

Sequence of the GAB15256 gene was obtained from the Transporter Protein Analysis Database (http://www.membranetransport.org/) (Paulsen *et al*., 1998: Ren *et al*., 2006). Restriction sites in the gene of interest was mapped using Webcutter 2 (https://users.unimi.it/camelot/tools/cut2.html), which allowed checking for the presence of any *EcoRI* or *PstI* restriction sites that were used for cloning. Primers were designed to introduce in-frame *EcoRI* site (GAATTC) at the 5` ends and *PstI* site (CTGCAG) at the 3` ends of the gene to allow subsequent ligation with *EcoRI-PstI* digested pTTQ18-His6 (Table 2.3). The forward and reverse primers were: 5`CCGGAATTCGCATATGACGGCGGC3` and 5` AAAACTGCAGCGCGGACCGGGGCATC 3` respectively. These primers were used to PCR for amplification of desired gene from genomic DNA of *Arthrobacter globiformis*.

Genes amplification from bacterial genomic DNA

PCR was performed using a Bio-rad thermocycler (Waltham, USA) with total reaction volumes of 50 µL by adding the reagents shown in Table 1 and using-conditions shown in Table 2. Plasmid DNA was isolated from bacterial overnight cultures and purified using a QIAprep Spin Miniprep Kit and protocol from QIAGEN Ltd., UK. The PCR product and plasmid pTTQ18 were cut with restriction enzymes *EcoRI* and *PstI*. Ligation of the gene of interest into plasmid DNA was catalysed by T4 DNA ligase followed by transformation.

After successful amplification, the DNA fragment was subsequently purified and the integrities of the cloned genes were recognized by automated DNA sequencing to ratify that each gene without mutation had been cloned and was inserted with a correct orientation into pTTQ18.

Table 1. PCR reaction components.

PCR reaction modules	**Final concentration**	**Final volume (**µL)
Sterile water	-	39
pfu Turbo buffer (10 x)	1 x	5
dNTPs (10 mM each)	0.2 mM	1
Genomic DNA (50 ng/µL)	50 ng	1
Forward primer (10 µM)	0.3 µM	1.5
Reverse primer (10 µM)	0.3 µM	1.5
pfu Turbo polymerase (100 units)	2.5 units	1
Total volume		**50**

Table 2. PCR reaction conditions.

Steps	**Temperature (^{o}C)**	**Time**	**Number of cycles**
Warm up	95	3 minutes	x 1
Denaturation	95	30 seconds	x 30
Annealing	58	30 seconds	
Extension	72	120 seconds	
Final extension	72	10 minutes	x 1
Hold	4	∞	x 1

Recombinant protein purification by immobilized metal affinity chromatography (IMAC)

Membranes from whole cells *E. coli* BL21(DE3) were prepared using the water lysis method (Witholt *et al*., 1976; Ward *et al*., 2000) and were resuspended in solubilisation buffer (20mMTris-HCl pH 8.0; 20 mM imidazole pH 8.0; 300 mMNaCl; 20% v/v glycerol; 1% w/v DDM) and mixed for 2 h at 4 ºC. The membranes were sedimented by spinning at 100,000 x g for 1 hour at 4ºC. The supernatant was incubated with Ni-NTA resin for 2 h at 4ºC with mixing. The recombinant proteins were purified by immobilized metal affinity chromatography followed by elution with a high concentration of imidazole buffer (10 mM Tris-HCl pH 8, 200 mM imidazole pH 8, 2.5% v/v glycerol, 0.05% DDM). The eluted samples were concentrated up to 3ml and applied to a BioRad Econo-pac 10 DG desalting column followed by wash buffer (10 mM Tris-HCl pH 7.6; 2.5% v/v glycerol; 0.05% DDM). The purified eluted fraction was concentrated to 5-20 mg/mL and stored at -80ºC.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE).

Proteins were resolved by SDS-PAGE using a 15% v/v resolving gel and a 4% v/v stacking gel. For gel analysis samples containing 15 µg of protein were solubilised in 4x sample loading buffer (60 mM Tris-HCl pH 7.2, 10% v/v glycerol, 2% w/v SDS, 0.005% bromophenol blue, 3% β-mercaptoethanol). Proteins were visualized by staining the gel in EZBlue stain (Wilson, 1979).

Circular dichroism spectroscopy

To determine secondary structure content of purified proteins, far-UV (180-260 nm) circular dichroism spectroscopy measurements were performed using a CHIRASCAN instrument (Applied Photophysics, UK) at 20 ºC with constant liquid nitrogen flushing. Samples were prepared in a Hellma quartz cuvette of 1.0 mm path length at a final protein concentration of 0.1 mg/mL in CD buffer (10 mM NaPi pH 7.5; 0.05% DDM). The thermal stability of protein was analysed by ramping the temperature from 5-90 ºC and finally back to 5 ºC. Spectra in the range of 180 to 260nm were recorded at 1nm intervals (Kōji *et al.*, 1994).

RESULTS AND DISCUSSION

The GAB15256 gene was amplified from *Arthrobacter globiformis* (Conn 1928) Bergey *et al*. 1930 genomic DNA with 5` and 3` cloning primers. PCR product of 1534bp was obtained (Fig. **1**) and cloned into pTTQ18 plasmid. The recombinant plasmid DNA was sequenced by Beckman Coulter Genomics (BCG, UK). The results verified that the gene GAB15256 and the recombinant plasmid pTTQ18 were constructed successfully. Also the nucleotide sequence of the cloned DNA fragment displayed 100% identity with the GAB15256 nucleotide sequence in NCBI using the multiple alignment tool Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). Bioinformatics analyses have reported this protein to be associated with the transport of myo-inositol which is necessary requirement of most of the organism. However, GAB15256 protein is a novel protein with unknown function and mechanism yet. After successful cloning the membranes containing amplified expression of GAB15256 were solubilised with 1% DDM and the His_6-tagged protein was purified by immobilised metal affinity chromatography using a Ni-NTA resin. Using the polypeptide sequence of GAB15256 (His_6) and the Expasy M_w calculator (http://web.expasy.org/compute_pi/), the theoretical mass of GAB15256 (His_6) was 56693.72Da and it migrates on SDS-PAGE at a position of 38.0 kDa which is a size difference of ~68%. This is consistent with SDS-PAGE analysis of other memberanes proteins, which travel at about 60-70% of their predicted sizes (Bettaney, 2008; Ma, 2010; Ma *et al*., 2016). It is widely recognized that membrane proteins migrate anomalously on SDS-PAGE gels at lower molecular weight positions than their actual molecular weights predicted from amino acid composition due to high binding of SDS, hydrophobic nature, or the secondary structure retention facilitating the migration through the gel (Rath and Deber, 2013; Ward *et al*., 2000). SDS-PAGE analysis confirms identity of GAB15256 protein where visible band was observed in Figure-2. Based on the purification shown in Fig. 2, the yield of purified protein was 1.2 mg per litre of cell culture. Lane 2 of SDS-PAGE contains the membranes solubilised with 1% DDM whereas lane 3 contains protein that did not bind to the Ni-NTA column (unbound fraction). Lane 4 shows 90 percent pure GAB15256 protein that was obtained with IMAC using a Ni-NTA resin which is one of the most efficient methods of protein purification. The NTA ligand contains four chelating sites that are interacting with metal ions. NTA usually occupies four of the ligand binding sites in the co-ordination sphere of the Ni^{+2} ion, leaving two sites to interact with the hexahistidine tag (Qiagen, 2003). Purified protein GAB15256 was eluted with buffer containing a high concentration of imidazole (200 mM); this dissociates the His_6-tagged proteins because they can no longer compete for binding sites on the Ni-NTA resin (Lane 4). The purity of the isolated protein was deemed suitable for subsequent biochemical and biophysical analysis. Generally, it is recommended that before performing biochemical and biophysical methods for characterization of protein, it is necessary to validate

the secondary structure of protein. CD spectroscopy is a valuable biophysical technique for examining the folding, conformational changes and determination of protein secondary structure (Wallace *et al*., 2003; Miles and Wallace. 2016). Therefore, far-UV CD spectroscopy (180-260 nm) was performed to measure the secondary structure content of the purified GAB15256 protein which was solubilised in DDM detergent. The resultant CD spectrum obtained of purified GAB15256 was indicative of a predominantly α-helical protein with a peak at ~192 nm and troughs at 209 nm and 222 nm (Fig. 3A). Importantly, no beta-sheet structure was revealed for recombinant GAB15256, these results were consistant with other reported studies (Wallace *et al*., 2003; Kelly *et al*., 2005; Bulheller *et al*., 2007; Bettaney *et al*., 2013). These measurements also established that the purified GAB15256 protein was correctly folded and had retained its alpha helical secondary structure after passing through various steps of the purification. This was an important control prior to further biophysical or biochemical analyses using the purified protein.

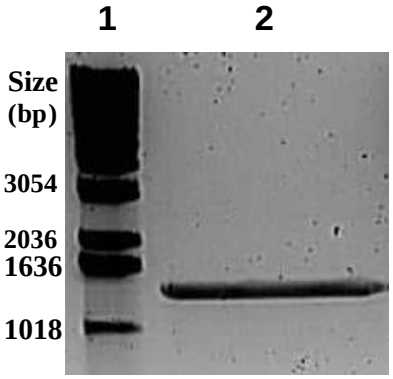

Fig. 1. PCR amplification of GAB15256 gene. PCR products was analysed on a 1% agarose gel. Lanes: 1 kb DNA ladder (left) and PCR products (right).

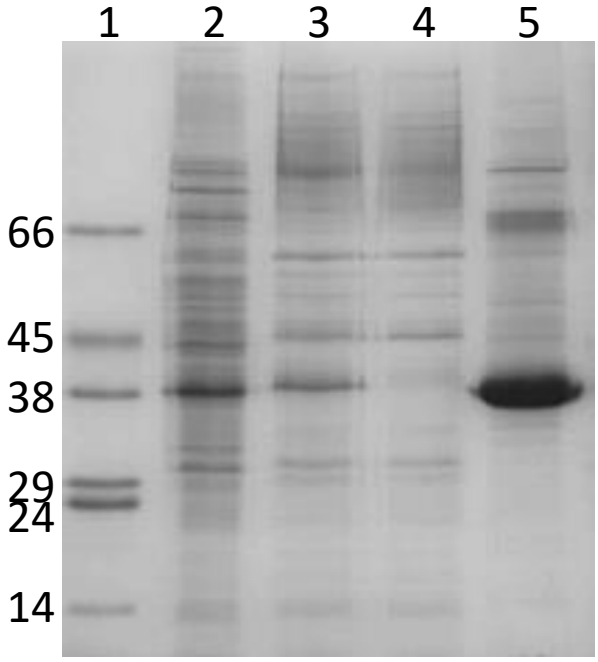

Fig. 2. Coomassie blue stained 15% SDS-PAGE showing the purification of GAB15256 protein from membrane. 15 µg samples loaded as follows: (1) molecular weight marker; (2) insoluble pellet; (3) supernatant membrane; (4) unbound flow through; (5) purified protein. Lanes 5 is of particular interest as they represent the purified fractions, respectively.

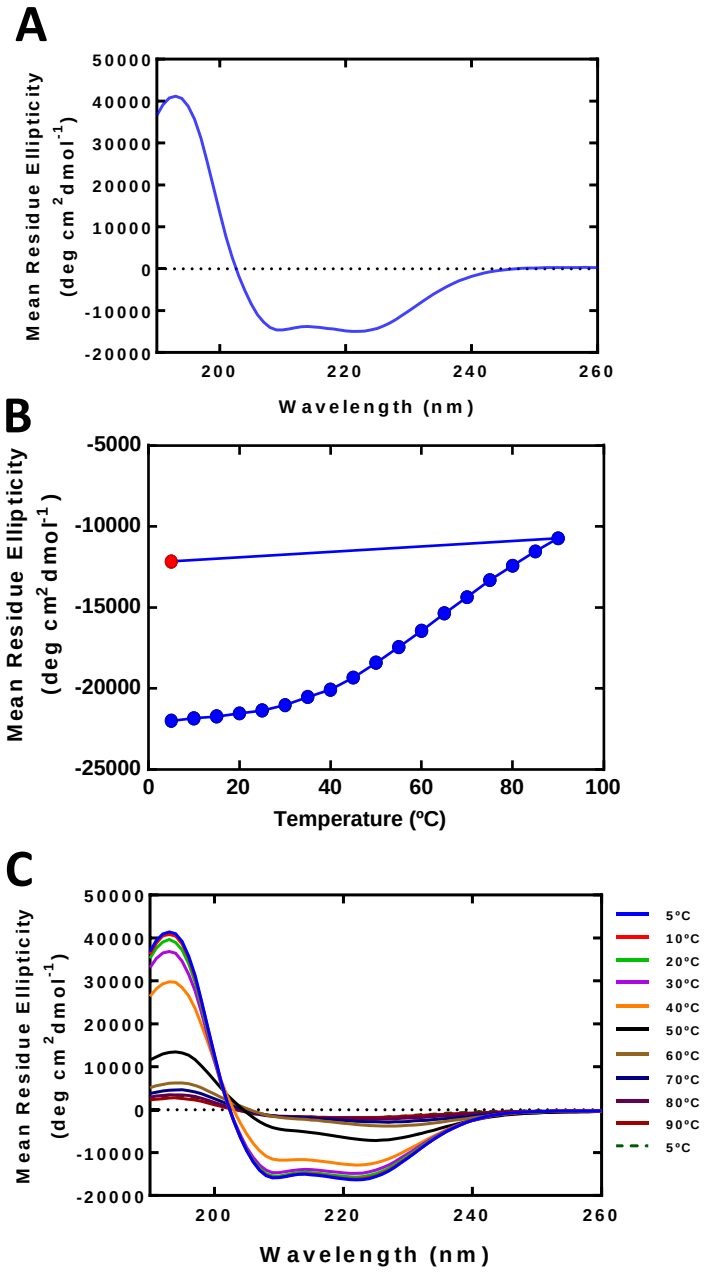

Fig. 3. Far-UV CD analysis of purified GAB15256 protein. (A) Far-UV (180-260 nm) CD spectrum for purified AAN69889 protein. Measurements were performed using a CHIRASCAN instrument (Applied Photophysics, UK) at 20 ºC with constant nitrogen flushing. Samples were prepared in a Hellma quartz cuvette of 1.0 mm pathlength at a final protein concentration of 0.15 mg/ml in CD buffer (10 mM NaPi pH 7.5; 0.05% DDM). (B) Thermal unfolding of AAN69889 proteinover the concentration range 5-90 ºC and finally back to 5 ºC monitored at a wavelength of 209 nm. (C) Full spectra showing thermal unfolding of AAN69889 protein over the given temperature range.

The thermal stability of GAB15256 was analysed by ramping the temperature from 5-90 ºC and finally back to 5 ºC (Fig. 3B & C). The final scan at 5 ºC shows whether the purified protein had refolded into its native secondary structure. GAB15256 loses it secondary structure with rising temperature and comparisons of spectra at 5 ºC prior to thermal denaturation and at 90 ºC following denaturation showed three main differences (Fig. 3C). Firstly, a decrease in the negative signals at both 209 nm and 222 nm and secondly significant reductions in the amplitude of the signal at 192 nm. These changes were indicative of reductions in the alpha helical content and therefore of

thermal denaturation. At 90 ºC there was still a positive signal at ~192 nm and negative signals at ~209 nm and ~222 nm suggesting that the protein was not fully unfolded. With increasing temperature the signal at 209 nm began to change significantly at ~40 ºC and a melting temperature of 46.2 ºC was estimated using Global Analysis CD software 3. The GAB15256 protein is therefore reasonably stable for performing biophysical assays using a temperature range of 18-25 ºC. On returning the temperature to 5 ºC, there was no evidence that suggests refolding of the protein, similar to reported results for other transport proteins (Bettaney, 2008; Ma, 2010; Sukumar, 2012; Jackson *et al*., 2013). This study established a series of protein purification steps for isolation of desired protein from a complex mixture and laid down the foundation for characterization of similar proteins.

Conflict of interest:
None declared.

REFERENCES

Berman, T. and B. Magasanik (1966). The pathway of myo-inositol degradation in Aerobacter aerogenes. Ring scission. *J. Biol. Chem*, 241: 807–813.

Bettaney, K. (2008). *Characterisation of membrane transport proteins of bacteria – from gene to crystal*. Ph.D. Thesis. University of Leeds, Leeds, UK.

Bettaney, K.E., P. Sukumar, R. Hussain, G.Siligardi, P.J. Henderson and S.G. Patching (2013). A systematic approach to the amplified expression, functional characterization and purification of inositol transporters from *Bacillus subtilis*. *Molecular membrane biology*, 30 (1): 3-14.

Boutte, C. C., B. S.J. A. Srinivasan, J. A. Flannick, A. F. Novak, A. T. Martens, S. Batzoglou, P. H. Viollier and S. Crosson (2008). Genetic and computational identification of a conserved bacterial metabolic module. *PLoS Genet*, 4 (2): 1-12.

Bulheller, B.M., A. Rodger and J. D. Hirst (2007). Circular and linear dichroism of proteins. *Physical Chemistry Chemical Physics*, 9 (17): 2020-2035.

Galbraith, M. P., S. F. Feng, J. Borneman, E. W. Triplett, F. J. de Bruijn and S. Rossbach (1998). A functional *myo*-inositol catabolism pathway is essential for rhizopine utilization by Sinorhizobium meliloti. *Microbiology*, 144: 2915–2924.

Jackson, S.M., S.G. Patching, E. Ivanova, K. Simmons, S. Weyand, T. Shimamura, F. Brueckner, S.I. Suzuki, S. Iwata and D.J. Sharples (2013). Mhp1, the Na^+-hydantoin membrane transport protein. *Encyclopedia of Biophysics.* Pp.1514-1521.

Jones, D. and R. M. Keddie (1992). *The procaryotes*. (In A. Balows, H. G. Tru¨per, M. Dworkin, W. Harder, and K. H.) Schleifer (ed.) Springer-Verlag, New York. The genus *Arthrobacter*. pp. 1283–1299.

Kawsar, H. I., K. Ohtani, K. Okumura, H. Hayashi and T. Shimizu (2004). Organization and transcriptional regulation of myo-inositol operon in *Clostridium perfringens*. *FEMS Microbiol. Lett*, 235: 289–295.

Kelly, S.M., T. J. Jess and N.C. Price (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics.*1751(2), 119-139.

Kōji Nakanishi; Nina Berova; Robert Woody (1994). Circular dichroism: principles and applications. VCH. p. 473

Krings E., K. Krumbach, B. Bathe, R. Kelle, V. F. Wendisch, H. Sahm and L. Eggeling (2006). Characterization of myo-inositol utilization by *Corynebacterium glutamicum*: the stimulon, identification of transporters, and influence on L-lysine formation. *J. Bacteriol.*, 188: 8054–8061.

Ma, P. (2010). *Structure-activity relationships of membrane proteins: The NCS-1 family of transports and sensor kinases of two-component systems.* Ph.D. Thesis, University of Leeds, Leeds, UK.

Ma, P., S.G. Patching, E. Ivanova, J. M. Baldwin, D. Sharples, S.A. Baldwin, P. J. Henderson (2016). Allantoin transport protein, PucI, from *Bacillus subtilis*: evolutionary relationships, amplified expression, activity and specificity. *Microbiology*, 162(5): 823-836.

Miles, A. J. and B. Wallace (2016). Circular dichroism spectroscopy of membrane proteins. *Chemical Society Reviews*,45(18): 4859-4872.

Mullaney, E. J., C. B. Daly and A. H. J. Ullah (2000). Advances in phytase research. *Adv. Appl. Microbiol*, 47: 157–199.

Parthasarathy, L. K., R. S. Seelan, C. Tobias, M. F. Casanova and R. N. Parthasarathy (2006). Mammalian inositol 3-phosphate synthase: its role in the biosynthesis of brain inositol and its clinical use as a psychoactive agent. *Subcellular Biochemistry*, 39: 293–314.

Paulsen, I.T., M. K. Sliwinski and M. H. Saier (1998). Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities. *Journal of molecular biology*, 277(3): 573-592.

Poole, P. S., A. Blyth, C. J. Reid and K. Walters (1994). Myo-Inositol catabolism and catabolite regulation in Rhizobium leguminosarum bv. viciae. *Microbiology*, 140: 2787–2795.

Qiagen (2003). The QIAexpressionist- A handbook for high level expression and purification of 6 x His-tagged proteins, QIAGEN Gmbh, QIAGEN Inc. and QIAGEN LTD.

Rath, A. and C.M. Deber (2013). Correction factors for membrane protein molecular weight readouts on sodium dodecyl sulfate–polyacrylamide gel electrophoresis. *Analytical biochemistr*, 434(1): 67-72.

Ren, Q., K. Chen, K. and I.T. Paulsen (2006). TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic acids research.* 35(suppl_1): D274-D279.

Roberts M. F. (2006). *Inositol in bacteria and archaea*, *In:* Majumder AL, Biswas BB, editors. (ed), *Biology of inositols and phosphoinositides*, vol 39 Springer, New York, NY, pp 103–104.

Robinson, A.S. 2011. Production of membrane proteins: strategies for expression and isolation. John Wiley & Sons.

Sukumar, P. (2012). *Molecular dissection of substrate and inhibitor binding to the D-Galactose-H^+ symport protein (GalP) from Escherichia coli–the bacterial homologue of GLUTT1.* Ph.D. Thesis, University of Leeds, Leeds, UK.

Tatusova, T., M. DiCuccio, A. Badretdin, V. Chetvernin, E. P. Nawrocki, L. Zaslavsky, A. Lomsadze, K. D. Pruitt, M. Borodovsky and J. Ostell (2016). Prokaryotic genome annotation pipeline. *Nucleic Acids Res,* 44(14): 6614-24.

Turner, B. L., M. J. Paphazy, P. M. Haygarth and I. D. McKelvie (2002). Inositol phosphates in the environment. *Philos. Trans. R. Soc. Lond. B Biol. Sci*, 357: 449–469.

Wallace, B., J. Lees, A.Orry, A. Lobley and R.W. Janes (2003). Analyses of circular dichroism spectra of membrane proteins. *Protein Science*, 12(4): 875-884.

Ward, A., N. Sanderson, J. O'Reilly, N. Rutherford, B. Poolman and P. Henderson (2000). *The Amplified Expression, Indentificaiton, Purification, Assay and Properties of Histidine-Tagged Bacterial Membrane Transport Proteins*. Blackwell Press: Oxford.

Wilson C.M. (1979). Studies and critique of Amido Black 10B, Coomassie Blue R, and Fast Green FCF as stains for proteins after polyacrylamide gel electrophoresis. *Anal Biochem*., 96 (2): 263–78.

Witholt, B., M. Boekhout, M. Brock, J. Kingma, H. van Heerikhuizen and L. de Leij (1976). An efficient and reproducible procedure for the formation of spheroplasts from variously grown *Escherichia coli*. *Analytical biochemistry*, 74(1): 160-170.

Yoshida, K. I., D. Aoyama, I. Ishio, T. Shibayama and Y. Fujita (1997). Organization and transcription of themyo-inositoloperoniol of *Bacillus subtilis*. *J. Bacteriol.*, 179: 4591–4598.

(*Accepted for publication May 2019*)