MONASH University

Structural studies of sensing domains of transducer-like proteins from *Campylobacter jejuni* and *Helicobacter pylori*

Mayra Alejandra Machuca Perez

M.Sc. Biomedical Sciences; B.Sc. Bacteriology and Clinical Laboratory

A thesis submitted for the degree of Doctor of Philosophy

At the Department of Microbiology, Monash University, Australia

December 2017

Copyright notice

Contents

List of figures

List of tables

Abstract

Chemotaxis and motility have been identified as important colonisation factors for both *Campylobacter jejuni* and *Helicobacter pylori.* Chemotaxis is the directed movement in response to extracellular chemical gradients, which allows motile bacteria to find optimum environmental conditions for their proliferation. Environmental stimuli are detected by a set of membrane-associated chemoreceptors also known as methyl-accepting chemotaxis proteins (MCP) or transducer-like proteins (Tlp). The signal molecules either interact directly or indirectly with the ligand binding domain (LBD) of chemoreceptors, which then transmit the information to cytoplasmic components that control flagellar rotation.

In the present study, a comprehensive structural analysis of the LBD from three chemoreceptors (Tlp1 and Tlp3 from *C. jejuni* and TlpC from *H. pylori*) was carried out. Analysis of the crystal structure of Tlp1, Tlp3 and TlpC LBDs revealed that they belong to the double Cache superfamily of chemoreceptors, which is the most abundant domain among extracellular sensors in prokaryotes. dCache domains are comprised of a membrane-distal and membrane-proximal subdomain, each of which harbour a putative ligand-binding pocket.

Tlp1-LBD was co-crystallised with aspartate; however, analysis of the high-resolution structure did not show any electron density corresponding to this amino acid bound to Tlp1. In agreement with the crystallographic analysis, no binding between Tlp1-LBD and aspartate was detected by isothermal titration calorimetry. Although previous studies reported Tlp1 as an aspartate receptor of *C. jejuni,* the structural and biophysical experiments performed in this study demonstrated that Tlp1 does not sense aspartate directly and suggests that chemoeffector recognition is likely to occur indirectly through interaction with an as-yet-unidentified binding protein.

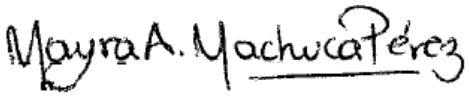
Tlp3 has been suggested as a chemoreceptor for multiple ligands that can directly sense a broad range of signal molecules that include both attractants and repellents. In order to understand the structural basis for ligand recognition by Tlp3, the crystal structures of Tlp3-LBD and its complex with an attractant (isoleucine) were studied in this dissertation. Detailed structural analysis revealed that Tlp3-LBD binds isoleucine through its membrane-distal subdomain. Interestingly, five of the residues interacting with the main chain of the ligand were found to be highly conserved in other dCache amino acid receptors. Mutagenesis experiments showed that these conserved residues are essential for isoleucine binding. These findings evince that the mechanism for amino acid sensing is highly conserved in dCache

chemoreceptors. Moreover, via a high-throughput screening of metabolic substrates, Tlp3-LBD was found to recognise other amino acids such as leucine, valine and α-amino-N-valeric acid, all of which bind to the membrane-distal subdomain of Tlp3-LBD in an almost identical mode to isoleucine binding. The outcomes of this research strongly suggest that direct ligand sensing by Tlp3 is specific and limited to branched-chain aliphatic amino acids and that this dCache LBD does not directly bind multiple and diverse ligands, as previously reported.

Finally, *H. pylori* TlpC-LBD structure and chemoeffector specificity were determined. TlpC-LBD crystal structure revealed a small molecule ligand bound at the putative binding pocket of the membrane-proximal subdomain, which was identified as lactate by mass spectrometry. Isothermal titration calorimetry and site-directed mutagenesis confirmed direct binding of lactate to TlpC-LBD and identified the amino acid residues required for ligand binding at the membrane-proximal subdomain. Additionally, it was determined that *H. pylori* is attracted to lactate which can be utilised as a carbon or energy source. TlpC is the first example of a chemoreceptor with dCache LBD that directly senses its chemoeffector through the membrane-proximal subdomain and suggests that dCache domain proteins can utilise both subdomains for ligand recognition.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature: 

Print Name: Mayra Alejandra Machuca-Perez

Date: 10-12-2017

Thesis including published works general declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes five original papers published in peer reviewed journals, and one unpublished publications. The core theme of the thesis is structural and functional characterisation of transducer-like proteins. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, working within the Department of Microbiology under the supervision of Associate Professor Anna Roujeinikova and Dr. Terry Kwok-Schuelein.

In the case of chapter two, three and four my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution	Co-author, Monash student
2	Cloning, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for multiple ligands (CcmL).	Published	Designed experiments, performed experiments, analysed data and manuscript editing (48%)	1. Yu C. Liu, Conception, designed & performed experiments (20%) 2. Beckham S. Data analysis, manuscript editing (5%) 3. Roujeinikova A. Conception, designed experiments, analysed data, manuscript editing (27%)	NO
	Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains.	Published	Designed experiments, performed experiments, analysed data and manuscript editing (41%)	1. Yu C. Liu, Conception, designed & performed experiments (24%) 2. Beckham S., Data analysis, manuscript editing (5%) 3. Gunzburgb M., Data analysis, manuscript editing (5%) 4. Roujeinikova A., Conception, designed experiments, analysed data, manuscript editing (25%)	NO
3	Cloning, expression, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter*	Published	Designed experiments, performed experiments, analysed data and manuscript editing (53%)	1. Yu C. Liu, Conception, designed & performed experiments (20%) 2.Roujeinikova A., Conception, designed experiments, analysed data, manuscript editing (27%)	NO

	chemoreceptor for aspartate A (CcaA).				
	The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni* Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition.	Published	Designed experiments, performed experiments, analysed data and manuscript editing (53%)	1. Yu C. Liu, Conception, designed & performed experiments (15%) 2. Beckham S., Data analysis, manuscript editing (5%) 3. Gunzburgb M., Data analysis, manuscript editing (5%) 4. Roujeinikova A., Conception, designed experiments, analysed data, manuscript editing (22%)	NO
4	*Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate	Published	Designed experiments, analysed data, manuscript editing 50%	1. Johnson K., Designed experiments, analysed data, manuscript editing (12%) 2. Yu C. Liu, Designed experiments, manuscript editing (12%) 3. Steer D., Data analysis, manuscript editing (3%) 4. Ottemann K., Designed experiments, analysed data, manuscript editing (7%) 5. Roujeinikova A., Conception, designed experiments, analysed data, manuscript editing (16%)	NO

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: Mayra A. Machuca Pérez **Date:** 10-12-2017

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: **Date:** 13-12-2017

List of publications during enrolment

Chapter two:

1. **Machuca, M.A.**, Liu, Y.C., Beckham, S.A., Roujeinikova A. Cloning, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for multiple ligands (CcmL). Acta Crystallogr. F Struct. Biol. Commun, 2015. 71 (Pt 2): p. 211-6.

2. Liu, Y.C., **Machuca, M.A.**, Beckham, S.A., Gunzburg, M.J., Roujeinikova, A. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. Acta Crystallogr. D Biol. Crystallogr., 2015. 71 (Pt 10): p. 2127-36.

Chapter three:

3. **Machuca, M.A.**, Liu, Y.C., Roujeinikova, A. Cloning, expression, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for aspartate A (CcaA). Acta Crystallogr. F Struct. Biol. Commun., 2015. 71 (Pt 1): p. 110-3.

4. **Machuca, M.A.**, Liu, Y.C., Beckham, S.A., Gunzburg, M.J., Roujeinikova, A. The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni* Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition. J. Struct. Biol., 2016. 194 (2): p. 205-213.

Chapter four:

5. **Machuca, M.A.**, Johnson, K.S., Liu, Y.C., Steer, D.L., Ottemann, K.M., Roujeinikova, A. *Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate. Sci. Rep., 2017. **7** (1): p. 14089.

Acknowledgements

This work would not have been possible without the support of several people that help in its completion. I would like to thank my supervisor Associate Professor Anna Roujeinikova, for her continuous support, guidance and encouragement throughout my PhD candidature. I am very grateful for accepting me as a student and helping me succeed in every challenge encountered during the development of this project. Thank you Anna for trusting me by giving me full autonomy, for your patience throughout my candidature and your valuable constructive criticism that has helped me improve as a researcher. I would like to thank my co-supervisor Dr. Terry Kwok-Schuelein and my PhD panel members Dr. Hans Netter, Associate Professor John Boyce and Dr. Matthew Belousoff for their valuable inputs and support during my candidature.

A very special thanks to Simone Beckman for all your invaluable knowledge, especially for the advice and training and helping me with the experimental difficulties I faced. Thank you Simone for your willingness to contribute to my PhD project. I am also grateful to Gabrielle Watson and John Liu who kindly trained me and helped me become independent in the lab.

I want to thank my labmates Alex Timorikova, Joyanta Modak, Mohammad Mizanur Rahman, Abu Ud-Din and Mohammad Firoz Khan for all the shared knowledge, discussion and most of all for being friendly and supportive all the time.

This journey would not have been the same without the support and love of my friends. A very special thank you to Jason Marroquin who from the beginning was by my side, always with a supportive, joyful and positive attitude towards our PhD adventure. I am grateful that I met you at the beginning of this journey, thank you for your friendship, your patience and everything you have thought me through all these years. To my dear friend Cara Nethercott, thank you for letting me be part of your life. Thank you Cara for the encouragement during this journey and making it more fun, I cannot imagine surviving this PhD without your support and unconditional friendship.

I would also like to thank Yolanda Carreno, Wilson Alberto, Mehdi Youssefi Matak, William Rodriguez and Jian (Emilia) Li for being such spectacular friends to me, all the adventures we shared were wonderful and unforgettable.

Finally and most important, I give thanks to the people who helped and supported me from a great distance, my beloved mom and dad. Thank you for your guidance and for encouraging me to follow my dreams, I am here because of you! To my beautiful sister and best friend Ximena, thank you for all your love and support. To my partner in life, Juan Carlos, I could not have done this without you. Thank you for being patient, supportive and for your unconditional love.

I would like to thank to the Departamento Administrativo de Ciencia, Tecnología e Innovación COLCIENCIAS for the scholarship "Convocatoria 568 Doctorado en el Exterior" and Monash University for the Monash Graduate Scholarship (MGS), Monash International Postgraduate Research Scholarship (MIPRS) and Faculty Postgraduate Excellence Award.

I would like to dedicate this thesis to the persons that laid the foundations of my life…

my loving grandparents Cheli, Hernancito, Victorita and Julito.

List of abbreviations and symbols

4HB	Four-helix bundle
Å	Angstrom
Cache	Calcium channel and chemotaxis
CD	Circular Dichroism
°C	Degrees Celsius
dCache	Double Cache
HAMP	Histidine kinase, adenyl cyclase, MCP protein and phosphatase
HBM	Helical bimodular
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilo-Dalton
LBD	Ligand Binding Domain
MAD	Multiple Wavelength Anomalous Dispersion
MW	Molecular Weight
MCP	Methyl-accepting Chemotaxis Protein
PEG	Polyethylene Glycol
PDB	Protein Data Bank
PAS	Per-Arnt-Sim
RMSD	Root Mean Square Deviation
RR	Response Regulator
SAD	Single Wavelength Anomalous Dispersion
sCache	Single Cache
TCA	Tricarboxylic acid
TM	Transmembrane
Tm	Thermal Melting Temperature
Tlp	Transducer-like protein
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

Chapter 1: Literature review

1.2. *Helicobacter pylori*

H. pylori is a spiral-shaped, Gram-negative, flagellated, and microaerophilic bacterium that colonises the human stomach. This pathogen was discovered by Barry Marshall and Robin Warren in 1984, who first isolated the Gram-negative bacterium from gastric biopsies [1] and in 2005 won the Nobel Prize in Medicine for their discovery. Currently, *H. pylori* is known as the predominant microorganism of the human stomach, and it is estimated to colonise more than 50% of the world's population [2-4]. The prevalence of *H. pylori* infection varies between regions and countries, with the highest prevalences observed in Africa (79.1%), Latin America and the Caribbean (63.4%), and Asia (54.7%) [4,5]. In contrast, North America (37.1%) and Oceania (24.4%) exhibit the lowest prevalence of *H. pylori* infection [4,5].

Colonisation of the gastric mucosa by *H. pylori* is acquired in early childhood, and it has been proposed that transmission occurs *via* human-to-human, oral-oral, faecal-oral, and gastro-oral routes [3,4,6]. However, the exact mechanisms involved in *H. pylori* transmission to hosts remain unclear. Several risk factors have been associated with *H. pylori* infection, including living in areas with poor sanitation, living in rural or overcrowded areas, lower educational level, and low socioeconomic status [3,4,7,8]. Once colonisation of the host has occurred, *H. pylori* persists in the human stomach throughout life [9], where it may cause a range of different diseases. Most of *H. pylori* carriers never present clinical symptoms [10]; however, local inflammation of the gastric mucosa (gastritis) is observed in all asymptomatic patients. The most common clinical manifestations of this infection include gastric and duodenal ulcer disease, gastric adenocarcinoma, and primary gastric mucosal-associated lymphoid tissue (MALT) lymphoma [11-13]. Since 1994, *H. pylori* has been acknowledged as a type I carcinogen [14], and it is now considered the most common etiologic agent of infection-related cancers, being the causative agent in up to 89% of gastric cancer cases and approximately 6.2% of all cancers worldwide [15]. *H. pylori* has also been associated to a lesser degree with other diseases, including atrophic gastritis, dyspepsia, iron deficiency anaemia, and idiopathic thrombocytopenia purpura [16,17].

1.2.1. *H. pylori* virulence factors

Several colonisation and virulence factors have been identified in *H pylori.* Some play important roles in stomach colonisation and the establishment of infection, while others are

strongly associated with increased risk of disease development. These include motility, urease, adhesins, lipopolysaccharide (LPS), and toxins.

1.2.1.1. Motility

As a first step in stomach colonisation, *H. pylori* must find its way from the gastric lumen through the mucus gel to finally reach the gastric epithelial surface, its colonisation site. To reach its natural niche, *H. pylori* passes through the mucus barrier using its flagellar motility and sensing chemical gradients. *H. pylori* cells are highly motile owing to a unipolar bundle of two to six flagella [18,19]. Each flagellum is about 3-5 μm in length and consists of three main structures: a basal body within the bacterial wall that contains proteins involved in flagellar rotation, a torsion hook, and a flagellar filament.

The importance of motility in stomach colonisation by *H. pylori* has been demonstrated using different non-motile mutants. For example, studies using aflagellated mutants showed that these strains were unable to colonise stomachs in animal models [20-22]. Similar results were obtained with non-motile mutants expressing wild-type flagella but lacking the motor protein MotB [23], which is essential for flagellar rotation. These results support the importance of functional flagella and the motility conferred on the organism for establishing persistent infection. In agreement with flagellar motility being an essential colonisation factor for *H. pylori*, failures in the well-regulated process of flagellar biogenesis have major repercussions on motility, and consequently on colonisation and infection of the gastric mucosa [23,24].

To find its way towards the mucous layer and gastric epithelium, *H. pylori* uses its chemotaxis signalling system, by which the bacterium senses different host attractants or repellents and moves toward or away from them, respectively. *H. pylori* senses amino acids, mucin, urea, bicarbonate, cholesterol, lactate, low pH, zinc, and nickel [25-29]. These external molecules are detected by methyl-accepting chemotaxis proteins (MCPs), which transduce the signal to cytoplasmic components that alter flagellar rotation. By sensing environmental cues, *H. pylori* can direct its swimming and find its preferred place for colonisation. Non-chemotactic mutants, for example, show strong defects in colonisation ability and are outcompeted in coinfection by wild-type strains [30-33]. Furthermore, only a minor proportion of non-chemotactic mutants were found in contact with the gastric cell surface [33,34]. Thus, chemotaxis plays a role in establishing colonisation of the deeper layers of the mucus and gastric epithelium by directing the bacterium to swim towards its natural niche.

1.2.1.2. Urease

Although *H. pylori* is not an acidophile, it is able to survive the acidic conditions of the gastric lumen by producing the urease enzyme [35]. Urease is considered one of the most important virulence factor in *H. pylori*. This enzyme hydrolyses urea to produce ammonia and carbamate, the latter of which is spontaneously converted to ammonia and carbonic acid. As a result, the periplasmic and cytoplasmic environment of *H. pylori* is buffered, maintaining the pH close to neutral, which allows the bacterium to escape gastric acidity [36,37].

The biosynthesis of urease is under the control of the seven-gene cluster *ureABIEFGH* [38]. The *ureA* and *ureB* genes encode the two structural subunits of the urease enzyme [38], while *ureE, ureF, ureG*, and *ureH* encode the chaperone proteins that assemble the catalytic site of the enzyme and incorporate nickel into the active sites [39,40]. Urease consists of a dodecameric complex of UreA and UreB subunits with twelve active sites [41], which must incorporate nickel ions for enzyme activation [41,42].

Concentrations of about 1–5 mM urea have been detected in the gastric juice [43]. Urea from the external environment is imported into the bacterial cytoplasm by a membrane protein encoded by the *ureI* gene [44]. UreI is a proton-gated channel regulated by changes in external pH. When the external pH is below 6.5, the UreI channel opens, allowing the rapid transport of urea into the *H. pylori* cytoplasm and maximal production of ammonia and carbamate from urease [44,45]. In contrast, at neutral pH, UreI is inactive, preventing lethal alkalinisation of the cytoplasm by excessive production of ammonia by *H. pylori* urease [44,45].

At neutral pH, *H. pylori* does not require urease for optimal growth; however, this enzyme is key to its survival under acidic conditions [46,47]. Several studies using urease and *ureI* mutants have shown that both functional proteins are essential for *H. pylori* survival in gastric mucosa [35,48,49]. In summary, *H. pylori* uses the urease system to maintain periplasmic and cytoplasmic pH near neutrality and overcome gastric acidity.

1.2.1.3. Adhesins

A very important step in establishing persistent infection is the adherence of *H. pylori* to the gastric epithelium. This close interaction with gastric cells protects the bacterium from host clearance mechanisms, such as gastric peristalsis, bulk liquid flow, and continuous shedding and replenishment of the mucus layer. Binding of the pathogen to the host tissues is mediated by a group of outer membrane proteins known as adhesins.

For instance, initial binding of *H. pylori* to the gastric epithelium is mediated by the blood group antigen-binding adhesin BabA, which adheres to human fucosylated Lewis b blood group antigen in the gastric epithelium [50,51]. BabA binding facilitates the attachment of the *H. pylori* type IV secretion system (T4SS) apparatus to host cells, and thus the delivery of other virulence factors [52]. The *H. pylori* adhesin SabA binds to sialyl-Lewis A antigen [51,53], which is expressed by gastric tissues in response to chronic inflammation and contributes to the chronicity of *H. pylori* infection [53,54].

The *H. pylori* adherence-associated lipoproteins A and B (AlpA and AlpB) are involved in bacterial adherence [55]; however, their receptors are unknown. *In vivo* studies on *alpA* and *alpB* isogenic mutants have revealed that these strains have reduced adherence capability compared to that of their wild-type counterparts [56]. Another outer membrane protein involved in *H. pylori* adherence to human gastric cells is the outer inflammatory protein (OipA) [57]. As with AlpA and AlpB, host-specific OipA receptors have not been identified. However, the interaction between OipA and host tissues has been shown to increase production of interleukin-8 (IL-8) [58], which leads to infiltration of neutrophils into gastric mucosa [58]. OipA expression has been associated with carcinogenesis [59], as shown in studies using *oipA* mutants, in which a significant decrease in tumorigenesis was observed in an infection model of Mongolian gerbils [60]. Thus, adhesion of *H. pylori* to the gastric epithelium is a necessary step in host colonisation and enables the pathogen to remain at the infectious site and to secrete toxins into the cells that eventually cause tissue damage and the progression to gastric ulceration.

1.2.1.4. Lipopolysaccharide (LPS)

The LPS of *H. pylori* shares a common architecture with that of other Gram-negative bacteria [61]; however, several differences in the LPS of this pathogen have been identified [62,63]. Typical LPS architecture comprises three main structures: lipid A, the core polysaccharide, and the O-antigen. Lipid A attaches the LPS to the outer membrane and is the toxic component of the LPS. In *H. pylori*, lipid A is under-phosphorylated and under-acylated [62,64]. For example, compared to lipid A in other Gram-negative bacteria such as *Escherichia coli*, lipid A in *H. pylori* has one less phosphate group, and of the six fatty acid chains present in *E. coli* lipid A, *H. pylori* lipid A has only four [62,64]. Furthermore, the fatty acid chains are longer than those found in *E. coli* lipid A [62]. Altogether, these differences make *H. pylori* lipid A less

toxic than *E. coli* lipid A [65,66]. The lower endotoxic and immunological activities of *H. pylori* lipid A allow the pathogen to produce a chronic infection of the host tissue [62,66].

Differences in *H. pylori* LPS have also been identified at the level of the O-antigen [63,66], which consists of repeating sugar units. In most *H. pylori* strains, the O-antigen repeats mimic the ABO Lewis blood group antigens Lewis x (Le^x) and Lewis y (Le^y) [63,67,68], while Le^a, Le^b, Le^c, sialyl-Le^x, and H-1 determinants are found to a lesser degree [69]. Expression of Lewis antigens by *H. pylori* has been implicated in immune response evasion [70]. Bacterial molecular mimicry of the ABO Lewis antigens allows *H. pylori* to escape from the host humoral response by preventing the production of antibodies against these antigens shared by the host and bacterium [71,72]. However, studies in animal models and humans have shown that the *H. pylori* O-antigen also contributes to gastric autoimmunity, which causes gastric tissue injury and gastric atrophy in chronic infections [73-75]. Finally, O-antigen seems to facilitate bacterial colonisation and adhesion, as shown in strains expressing Le^x O-antigen, which binds the host galactoside-binding lectin, galectin-3 [76], and thus mediates adherence to the human antral gastric mucosa [77].

1.2.1.5. Toxins

Several virulence factors associated with gastrointestinal diseases have been identified In *H. pylori*. These include the secretory toxins, namely, vacuolating cytotoxin A (VacA) and cytotoxin-associated gene A (CagA) [78]. VacA is an intracellular-acting exotoxin that binds multiple host receptors, including receptor-like tyrosine phosphatases (RPTP-α and RPTP-β) [79,80], sphingomyelin [81], epidermal growth factor receptor [82], fibronectin, heparin sulphate [83], low-density lipoprotein receptor-related protein-1 [84], and glycosylphosphatidylinositol (GPI)-anchored proteins [85]. After binding to the host cell, VacA triggers its internalisation into the cell through endocytosis [86], and it is then localised into late endosomes and lysosomes [87]. Inside the host cells, this exotoxin induces formation of large cellular vacuoles [88]. Vacuolation affects the proteolytic activity of the endocytic pathway, including antigen processing, leading to failure of the immune response [89]. It also compromises protein transport from the trans-Golgi network to the late endosome, causing release of acid hydrolases and subsequent cellular damage [90,91]. VacA also acts as a pore-forming toxin, leading to the release of molecules contained inside gastric cells that are essential for *H. pylori* survival *in vivo*, including Fe^{3+}, Ni^{2+}, sugars, and several amino acids [92,93]. In this way, VacA facilitates nutrient acquisition and therefore *H. pylori* colonisation and growth in the gastric epithelium [92,94].

Another very important *H. pylori* exotoxin is CagA. The gene encoding CagA is part of the 40 kb locus known as the Cag pathogenicity island (Cag PAI) [95]. The Cag PAI encodes all genes required for the structure and assembly of the T4SS apparatus [96]. The *H. pylori* T4SS apparatus translocates CagA into human gastric epithelial cells [97]. Once inside the host cells, CagA localises in the cell membrane underneath the point of bacterial attachment [98]. CagA is then phosphorylated and interacts with multiple host proteins, including tyrosine phosphatase SHP-2 [99,100], E-cadherin [101], and Par1b/MAPK2 [100,102]. CagA interaction with host proteins leads to leakiness of tight junctions [98] and cell elongation, which are typical of *H. pylori*-infected cells [78]. Several studies have demonstrated that CagA and VacA cause direct cellular damage that results in gastric ulceration and subsequent development of gastric cancer [60,78,103].

1.3. *Campylobacter jejuni*

C. jejuni is a Gram-negative, microaerophilic, spirally curved, slender bacterium that is highly motile by means of bi-polar unsheathed flagella [104]. *C. jejuni* is frequently found colonising the gastrointestinal tract of wild and domesticated animals, resulting in a commensal relationship between the bacterium and the host [105-107]. Although it is a commensal organism in the intestinal tracts of these animals, *C. jejuni* infection in humans produces bacterial gastroenteritis known as campylobacteriosis. Human infection commonly occurs from consumption or handling of undercooked, contaminated poultry [108,109]. However, other significant sources of *C. jejuni* infection include consumption of unpasteurised milk and untreated water [110].

C. jejuni was first isolated from the blood and faeces of an individual with acute enteritis by Butzler and Skirrow in the earlier 70's [111,112], and it is now known as the leading cause of bacterial gastroenteritis worldwide [113-115]. The annual incidence of campylobacteriosis cases varies among different countries, with an increasing trend in several developed countries [115]. For example, in United States of America, Canada, and Australia, the latest reported annual incidence rates were 14, 35, and 110 cases per 100,000 population, respectively [114,116,117]. In developing countries in Africa, Latin America, Asia, and the Middle East, *C. jejuni* infections are endemic, mostly affecting young children [115,118].

The clinical manifestation of campylobacteriosis ranges from asymptomatic to acute gastroenteritis characterised by bloody diarrhoea, high fever, acute abdominal pain, nausea,

and general malaise [108,119]. *C. jejuni* infection is self-limiting, resolving within a week without any complications. However, pregnant women, the elderly, and patients with acquired immune deficiency syndrome (AIDS) are considered at higher risk, as campylobacteriosis can be chronic and more invasive in such individuals [108]. Additionally, *C. jejuni* causes severe sequelae, such as reactive arthritis, peripheral neuropathies, and Miller-Fisher and Guillain-Barrè syndromes in some patients [120,121].

1.3.1. *C. jejuni* virulence factors

C. *jejuni* expresses several virulence factors that enable it to successfully colonise the host intestine and overcome host immune response. Some of these virulence factors include motility, the cell capsule, lipooligosaccharide, adhesins, and toxins.

1.3.1.1. Motility

To establish colonisation and infection, *C. jejuni* must first avoid being cleared from the intestine by the motion of peristalsis, and it must penetrate the mucus barrier [110]. To accomplish this, *C. jejuni* uses motility conferred by its polar flagella [110]. In fact, the motility of *C. jejuni* has been observed to increase in highly viscous solutions [122], which might facilitate its swimming through the thick mucus lining the intestinal tract epithelium [122].

Studies using aflagellate mutants have shown that a fully motile and intact flagellum is required for intestinal colonisation of animals and humans [123-125]. In addition, non-motile mutants are unable to adhere to and subsequently invade intestinal cells [126,127]. These results showed that motility is not only required for colonisation of the host but is also essential for invasion of host tissue [123,126].

C. jejuni motile behaviour is controlled by chemotaxis. The *C. jejuni* genome contains fifteen different MCPs, or chemoreceptors, of which ten might be involved in the sensing of environmental signals. *C. jejuni* responds chemotactically to several compounds, such as amino acids, carbohydrates, organic acids, and constituents of bile and mucus, which are frequently found in the host intestine [128-130]. The importance of chemotaxis in *C. jejuni* pathogenesis was demonstrated in experiments using non-chemotactic mutants that failed to colonise the gastrointestinal tract of mice and chickens [131,132] and that were also shown to lose invasiveness *in vitro* [132,133].

In addition to providing motility, *C. jejuni* flagella also function as a type III secretion system (T3SS) for the export of proteins that control bacteria-host interactions [134-137]. The

proteins secreted through the flagella are known as *Campylobacter* invasion antigens (CiaA-H) and flagellar co-expressed determinants (FedA-D) [134,138,139]. *cia*$^-$ and *fed*$^-$ *C. jejuni* mutants exhibit diminished ability to invade and colonise the intestines of chickens [134,135,139,140].

1.3.1.2. Cell Capsule

C. jejuni expresses a polysaccharide capsule (CPS) that has been shown to contribute to its pathogenesis [110,141]. The CPS is the major serodeterminant of the Penner serotyping method of *C. jejuni* strains, which includes more than 60 serotypes [142,143]. The CPS is encoded by a genetic locus comprising over 30 genes involved in biosynthesis, polymerisation, and translocation of the CPS [142]. The CPS locus is one of the more variable loci among *C. jejuni* strains, and its variability is reflected in the diversity of polysaccharide capsular structures observed in *C. jejuni* [142]. Moreover, the gene expression of this region is phase-variable, resulting in the production of modified CPSs [142].

In addition to the variations in polysaccharide composition, the CPS of *C. jejuni* undergoes modifications by ethanolamine, glycerol, and O-methyl phosphoramidate groups [142]. Interestingly, some of these modifications modulate cytokine release and are important virulence determinants, as demonstrated by *in vivo* and *in vitro* models [144,145]. *In vitro* studies using mutants deficient in CPS production have revealed that the ability of these strains to adhere to and invade intestinal epithelial cells is compromised [143]. Furthermore, capsuleless mutants exhibit decreased resistance to human serum and exhibit deficient ability to colonise ferrets [143].

1.3.1.3. Lipooligosaccharide (LOS)

LOS is considered an important virulence factor for *C. jejuni* [110,141] and constitutes the major chemical component of its outer membrane [146]. Structurally, *C. jejuni* LOS is similar to the lipopolysaccharides (LPS) of other Gram-negative bacteria; however, it lacks the O-antigen portion consisting of polysaccharide subunit repeats [147]. The LOS structure of *C. jejuni* has been well-characterised and comprises lipid A, which links the LOS to the outer membrane, and the core [147].

The LOS core polysaccharide of *C. jejuni* strains is highly diverse in monosaccharide linkage and composition [141]. Interestingly, the terminal sugars of the core polysaccharide are modified with endogenously synthesised sialic acid, which makes them similar to carbohydrate epitopes of gangliosides found in nerve cells [147]. The mimicry of ganglioside structures may be a mechanism used by *C. jejuni* to evade the host immune response, since

mutants lacking terminal sialyl sugars in the core show increased sensitivity to human serum and immunogenicity [148]. Recognition of these outer membrane components by the immune response can produce cross-reactive antibodies that target the gangliosides, resulting in nerve damage and demyelinating neuropathies such as Guillian-Barrè and Miller Fisher syndromes [121].

1.3.1.4. Adhesins

Once *C. jejuni* has entered the human gastrointestinal tract and breached the mucus barrier, the pathogen adheres to the underlying epithelial cells of the intestine. Several *C. jejuni* adhesion factors, or adhesins, and their corresponding receptors expressed by the host cells have been identified. Some of the major known *C. jejuni* adhesins include *Campylobacter* adhesion to fibronectin (CadF) [149], fibronectin-like protein A (FlpA) [150], Jejuni lipoprotein A (JlpA) [151,152], and periplasmic binding protein 1 (PEB1) [153].

CadF and FlpA are two outer membrane proteins that bind to the extracellular matrix protein fibronectin on the host cells and cooperatively mediate *C. jejuni* attachment to intestinal epithelial cells [138,149,150,154]. Moreover, CadF has been shown to mediate the internalisation of *C. jejuni* during the invasion of host cells [155].

JlpA lipoprotein mediates adherence to intestinal epithelial cells by interacting with the surface-exposed heat shock protein Hsp90a [151]. *C. jejuni* mutants lacking *jlpA* show decreased adhesion to cultured intestinal epithelial cells [151,152]. Similar results were observed with *cagF* mutants, which presented lower levels of adherence to and invasion of intestinal epithelial cells and had a severe colonisation defect [156]. Another adhesin, PEB1, has been shown to play a key role in adherence to and invasion of epithelial cells *in vitro* and in intestinal colonisation in an *in vivo* model [157]. In addition to its adhesion function, PEB1 also acts as an aspartate/glutamate-binding protein and is essential for optimal growth in media with dicarboxylic acid as the sole carbon source [158].

1.3.1.5. Toxins

Some strains of *C. jeuni* produce cytolethal distending toxin (CDT), a membrane-associated protein that has been identified as a virulence factor required for the pathogenesis of the bacterium [159,160]. CDT is composed of three subunits, namely, CdtA, CdtB, and CdtC, which form an active complex [161]. CdtB is the active subunit that functions as a DNase, while CdtA and CdtC bind to the host cell [161,162]. CDT toxin stimulates the production of proinflammatory cytokine IL-8 and induces cell cycle arrest, cytoplasm distention, chromatic

fragmentation, and cell death [163,164]. The exact role of CDT in diarrhoeal disease caused by *C. jejuni* remains unclear. *In vivo* studies have shown that *cdtB* mutants exhibit decreased invasiveness in blood, liver tissue, and spleen tissue in a mouse model [165]. However, *C. jejuni* isolates lacking CDT are still able to cause campylobacteriosis in humans [166].

1.4. Chemotaxis

Chemotaxis is a mechanism by which bacteria control flagellar motility. This process allows bacteria to detect different signals (chemoeffectors) and find favourable environments for growth by moving, for example, towards beneficial compounds (attractants) and away from unfavourable compounds (repellents). As described previously, chemotaxis has been shown to be an important colonisation and virulence factor for both *H. pylori* and *C. jejuni*, as well as for many other pathogens [167-170]. In this process, chemoeffectors are sensed by a set of bacterial chemoreceptors, which are typically integral membrane proteins harbouring an extracytoplasmic ligand-binding domain (LBD). These chemoreceptors can be methylated at specific residues on their cytoplasmic portion, and are therefore termed as **m**ethyl-accepting **c**hemotaxis **p**roteins (MCPs) [171,172]. Once the stimuli are recognised, the signals are transduced by the MCPs to downstream proteins, which ultimately regulate flagellar motion.

1.4.1. Chemosensory signal transduction: The *E. coli* model

The chemotaxis process has been extensively investigated in the enteric pathogen *E. coli*, and it has been used as an example for understanding and comparing chemotaxis in other bacteria species [171]. In the *E. coli* system, there are five kinds of MCPs which exhibit different ligand specificities: the aspartate receptor, Tar; the serine receptor, Tsr; the ribose/galactose receptor, Trg; the dipeptide receptor, Tap; and the aerotaxis sensor, Aer [173-175] . Changes in the concentration of chemotactic ligands are detected by the LBD of these MCPs, after which the information is transmitted to the downstream signal transduction proteins.

The core unit of the chemosensory signal pathway in bacteria involves protein members of a two-component regulator family that includes the chemotaxis proteins CheW, CheA, CheY, and CheZ [171,176]. CheW is a scaffold protein with two docking domains, one of which binds to the MCP and the other to CheA, forming a ternary complex [171,176]. The histidine kinase CheA has autophosphorylation activity modulated by the sensing of a stimulus. For example, when repellents bind to MCP or in the absence of attractants, the autophosphorylation activity of CheA is activated [172]. Phosphorylated CheA binds and phosphorylates the response

regulator (RR) protein CheY (CheY-P). CheY-P interacts with the flagellar switch and induces a clockwise rotation of one or several flagella, which destabilises the flagella bundle, causing the bacterium to perform a tumbling movement [171,172]. Finally, the signal response in *E. coli* is terminated by the phosphatase CheZ, which interacts with CheY-P, accelerating its dephosphorylation [177]. In contrast to repellents, attractants binding to MCP inhibit the CheA kinase activity; thus, CheY is not phosphorylated and CheY-P levels diminish. Non-phosphorylated CheY does not interact with the flagellar switch, resulting in a counterclockwise rotation or swimming (run) motion of the bacterium [176].

Following the initial response, an adaptation occurs that enables the bacteria to go back to a pre-stimulus status and to sense new changes in chemoeffector concentration. Adaptation is achieved by methylation and demethylation of glutamate residues at the cytoplasmic signalling domains of the chemoreceptors [171,178]. In *E. coli*, two enzymes are involved in chemoreceptor adaptation: the methylesterase CheB and the methyltransferase CheR. CheR incorporates methyl groups onto specific glutamate residues of the MCP, while CheB removes these methyl groups. Similarly to CheY, CheB also interacts with CheA and is phosphorylated, which activates its methylesterase activity [171].

1.4.2. Chemosensory signal transduction in *H. pylori* and *C. jejuni*

Both *H. pylori* and *C. jejuni* have the same core unit of chemosensory molecules as does *E. coli,* with the presence of a CheA kinase, a CheW coupling protein, and the RR CheY. However, some differences can be highlighted in the chemotaxis signalling pathway of these two pathogens. Firstly, the CheA kinases of *C. jejuni* and *H. pylori* differ from canonical CheA proteins. In these two bacteria and most members of the *Epsilonproteobacteria*, CheA proteins are fused with an RR (response regulator) domain at their C-terminus, which is similar to the RR domain of CheY [167]. Although the exact function of the RR domain in CheA is unknown, it has been suggested that is aids dephosphorylation of CheY, acting as a "phosphate sink" [167,179].

Another interesting difference is the presence of additional chemotaxis proteins in *C. jejuni* and *H. pylori*, such as the coupling protein CheV [179,180]. CheV is a protein comprised of a CheW domain and an RR domain. In addition to CheW, *C. jejuni* and other *Epsilonproteobacteria* express one CheV protein [179], whereas *H. pylori* expresses three: CheV1, CheV2, and CheV3 [167,180]. *C. jejuni* CheV and *H. pylori* CheV3 share 41% sequence

identity. *C. jejuni* CheV and the three *H. pylori* CheVs play a role in chemotaxis and promote chemoreceptor–kinase interactions [129,181-183]; however, whether chemoreceptors prefer CheV over CheW for coupling with CheA remains unclear [167,183].

As previously mentioned, chemoreceptor adaptation is mediated by the CheB and CheR enzymes in *E. coli*. CheB is regulated by the CheA kinase, which phosphorylates CheB at its RR domain and thus activates its methylesterase activity. Although these two enzymes are expressed in *C. jejuni*, CheB lacks the RR domain and contains only the methylesterase domain [179,184]. With the absence of an RR domain in *C. jejuni* CheB, other mechanisms for regulating CheR and CheB function and for adaptation might be employed by *C. jejuni* [179]. Interestingly, chemoreceptors in *H. pylori* do not use methylation and demethylation as an adaptation mechanism, as both the CheB and CheR enzymes and the conserved methylation sites on chemoreceptors are absent from this pathogen [167,185]. Further investigation is required to fully elucidate the adaptation mechanisms employed by chemoreceptors in *C. jejuni* and *H. pylori.*

1.4.3. Chemoreceptors

Chemotactic signals are sensed by a set of chemoreceptors, also known as MCPs or transducer-like proteins (Tlps) [186]. Typically, MCPs consist of an LBD, a transmembrane portion, and a cytoplasmic signalling domain. The latter is a conserved structure among different bacteria and archaea species and harbours a HAMP (histidine kinase, adenyl cyclase, MCP and phosphatase) region, conserved methylation sites, and a signalling subdomain [175,187].

1.4.3.1. Chemoreceptor topology

Chemoreceptors can be membrane-anchored proteins with either intracellular or extracytoplasmic (periplasmic) LBDs or soluble receptors in the cytosol [188,189]. Thus, based on their membrane topology, MCPs are grouped into seven classes: I (Ia and Ib), II, III (IIIm and IIIc), and IV (IVa and IVb) (Figure 1.1)[188,190,191]. Class I (Ia and Ib) is a predominant MCP type in both bacteria and archaea [188]. Chemoreceptors with this topology harbour an extracytoplasmic LBD, a transmembrane (TM) domain, and a cytoplasmic signalling domain [188] (Figure 1.1). According to the number of TM helices, class I MCPs are subdivided into Ia (two TM helices) and Ib (one TM helix) subclasses. Subclass Ia consists of an extracytoplasmic LBD flanked by an N-terminal and a second TM helix, as well as a C-terminal cytoplasmic signalling domain. In contrast, subclass Ib MCPs lack the N-terminal TM helix [192]. Subclass Ia

has been further subdivided into two clusters, I (~150 amino acids) and II (~250 amino acids) [188], according to the size of their LBDs (Figure 1.1).

Class II chemoreceptors are membrane-embedded proteins with cytoplasmic LBDs and consist of an N-terminal LBD, followed by two TM helices and a cytoplasmic signalling domain. This MCP type is less common than other chemoreceptor topologies and seems to be restricted to bacteria [188]. Similarly to class II, class III chemoreceptors are also membrane-embedded MCPs comprised of 1 or up to 8 TM helices, and their LBDs can be located either within the membrane portion (subclass IIIm) or after the last TM helix at the cytoplasm (subclass IIIc)[188]. Finally, class IV MCPs are soluble cytoplasmic receptors, some of which have an identifiable LBD (IVa). This topology is more frequently found in archaea than in bacteria [188].

1.4.3.2. Chemoreceptor LBD diversity

In addition to having different topologies, chemoreceptors vary greatly in the amino acid sequence and domain architecture of their LBDs. To date, at least 80 different LBD types have been identified [188,193]; structures characterised as 4-helix bundle (4HB), helical biomodular (HBM), Per-ARNT-Sim (PAS), and calcium channel and chemotaxis (Cache) domains are the most abundant families (Figure 1.2).

4-helix bundle (4HB) domains

4HB domains are the most abundant LBD type among chemoreceptors [193]. The 4HB fold is present in chemoreceptors with class Ia membrane topology, in which the LBD is located in the extracytoplasmic space (periplasm) (Figure 1.2). This domain has been well-characterised in the Tar and Tsr chemoreceptors from *E. coli*. Tar mediates chemoattractant responses to aspartate and maltose by two different mechanisms [194-196]. Aspartate recognition occurs through direct binding of the amino acid to the 4HB LBD of the receptor [197,198], while maltose sensing occurs *via* recognition of a maltose-binding protein [196]. For ligand sensing, LBDs with the 4HB fold require dimerisation, as the ligand-binding site is located at the dimer interface and comprised of amino acid residues from both monomers [197]. Other chemoreceptors with a predicted 4HB LBD are *Pseudomonas aeruginosa* CtpH [199] and *Pseudomonas putida* PcaY [200] receptors, which directly recognise inorganic phosphate and a range of carboxylic acids, respectively.

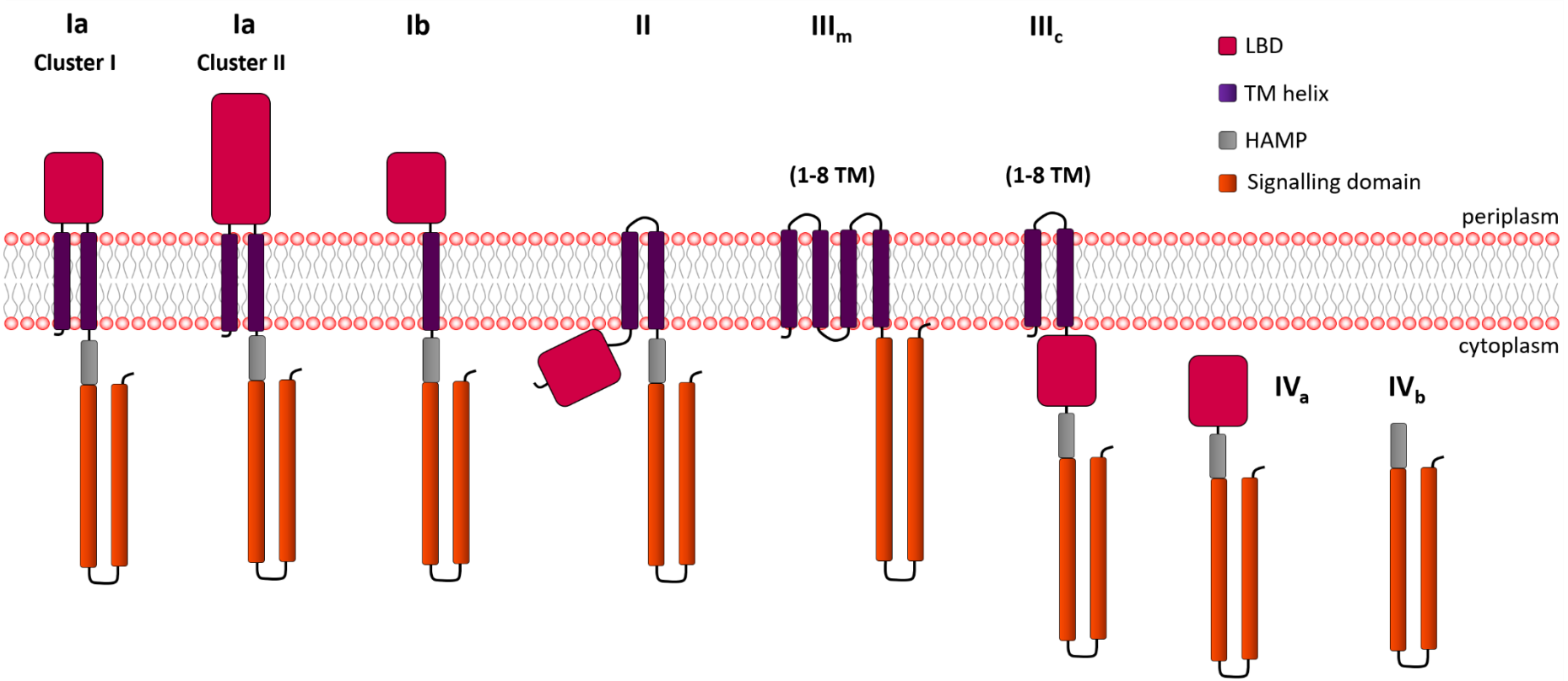


Figure 1.1. Topology diversity of MCPs. Classification of chemoreceptors based on membrane topology and LBD. The different classes are labelled. The larger LBD shown for the Ia cluster II receptor represents its longer amino acid sequence.

Per-ARNT-Sim (PAS) domains

PAS domains are widespread in all kingdoms of life [201] and are involved in the regulation of diverse processes [202-204]. Chemoreceptors with PAS domains are abundant in bacteria and archaea [188,193]. A PAS module comprises a single globular fold with a central antiparallel β-sheet with five strands and several α helices [204]. This domain is presented exclusively in chemoreceptors with cytoplasmic LBDs [188,205]. *E. coli* Aer receptor is an example of PAS-containing LBD with a class II membrane topology [206]. The Aer PAS domain binds flavin adenine dinucleotide (FAD) cofactor and mediates chemotaxis towards oxygen and favourable environments for energy production [206,207]. Another MCP with a PAS LBD is the *P. aeruginosa* Aer2/McpB, which binds a heme group and recognises O_2, NO, CO, and cyanide [208,209].

Calcium channel and chemotaxis (Cache) domains

Initially, Cache domains were annotated as PAS domains based on three-dimensional structural similarities. However, it has become clear that PAS domains are exclusively of cytosolic location, and that the "PAS-like" structures observed in extracellular LBDs belong to the Cache superfamily [205]. Cache is the most abundant LBD type in prokaryotes [188,193], and together with the 4HB and PAS superfamilies, accounts for more than 80% of all chemoreceptors with identifiable LBDs [193].

Two subgroups of LBD have been identified in the Cache superfamily: the single Cache (sCache) and double Cache (dCache) domains. One structurally characterised sCache is the TlpB class Ia (cluster I) receptor of *H. pylori*. The TlpB 3D-structure comprises an N-terminal long α-helix and three predicted strands that form an antiparallel sheet similar to that present in the core of the PAS domain structure [204,210] (Figure 1.2). The TlpB chemoreceptor mediates the *H. pylori* repulsion response to acidic conditions [210,211] and the AI-2 quorum sensing molecule [212], as well as chemoattraction to urea [210,213]. Urea binds to TlpB with high affinity [210], while AI-2 recognition appears to involve an indirect binding mechanism by interaction with periplasmic binding proteins [214]. Another chemoreceptor with an sCache LBD is the *P. putida* KT2440 McpP receptor, which directly binds several carboxylic acids, such as acetate, pyruvate, propionate, and L-lactate [215].

The dCache is more common than the sCache domain and is the most ubiquitous extracellular LBD in prokaryotes [188,193]. This domain has a bimodular structure composed of a long stalk helix at the N-terminus and two Cache modules (PAS-like domains) that are

arranged in tandem (membrane-proximal and membrane-distal), each of which harbour a putative ligand-binding pocket [216,217] (Figure 1.2). To date, several chemoreceptors with dCache LBDs have been investigated, including the cluster II chemoreceptors Mlp24 from *Vibrio cholerae* [216], Mlp37 chemoreceptor from *Vibrio parahaemolyticus* [217], PctA from *P. aeruginosa* [218], and McpB and McpC from *B. subtilis* [219,220]. In contrast to 4HB modules, dimerisation is not required for ligand binding to the dCache LBD, as the ligand-binding site is located within the Cache module putative pocket [217,218].

dCache LBDs recognise diverse signals using either direct or indirect mechanisms. Ligands that directly interact with the dCache LBD include amino acids [217,219,221-223], organic compounds [29,217], pyrimidines [224], and purines [225]. In most characterised dCache LBDs, direct sensing involves binding of the signal molecule to the putative binding pocket at the membrane-distal module [216,217,219-221]. However, with the presence of a putative binding pocket at the membrane-proximal subdomain, dCache modules could theoretically employ the membrane-distal subdomain, the membrane-proximal subdomain, or both for ligand recognition [29,226].

Helical bimodular (HBM) domains

Another LBD with two modules is the HBM domain. The McpS chemoreceptor of *P. putida* KT2440 was the first receptor identified with a HBM LBD [227]. McpS belongs to the Ia subclass of receptors with a cluster II LBD. The 3D structure of McpS-LBD shows that is composed of two long and four short helices arranged into two helical modules [227] (Figure 1.2). Like dCache domains, each module of the HBM fold contains a ligand-binding site. The McpS LBD directly binds the tricarboxylic acid cycle intermediates malate and succinate at its membrane-proximal module, whereas acetate binds to the membrane-distal module [227]. Interestingly, ligand binding to both modules stimulates McpS additively [227]. It has been suggested that this family of chemoreceptors primarily recognises organic acids [228]. Other receptors from this family are the *P. putida* KT2440 McpQ receptor [229], *P. fluorescens* Pf0-1 McpS receptor [230], and *P. putida* F1 McfS receptor [231], all of which are involved in chemotaxis to organic acids.

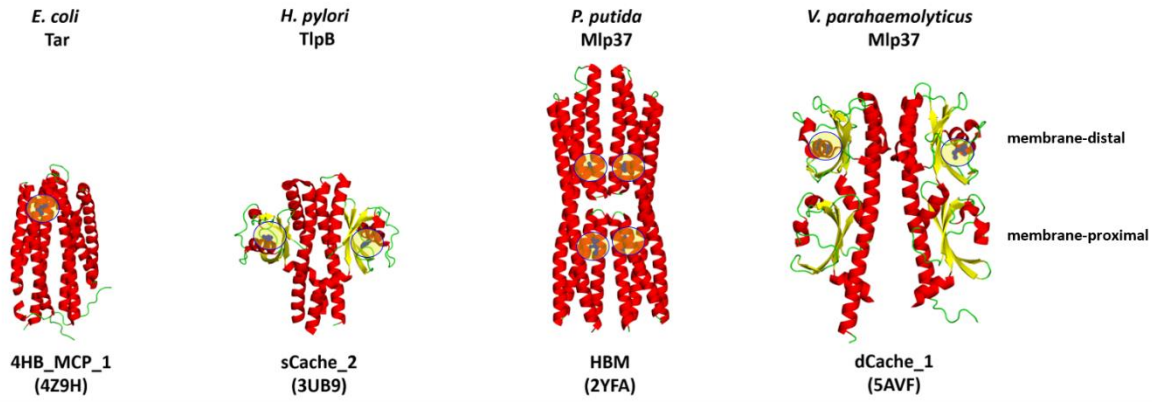


Figure 1.2. Structures of four of the most abundant LBD types. Structures of LBD dimers for *E. coli* Tar (in complex with aspartate) [198], *H. pylori* TlpB (in complex with urea) [210], *P. putida* McpS (in complex with malate) [227], and *V. parahaemolyticus* Mlp37 (in complex with taurine) [217]. LBDs are coloured by secondary structure and the bound ligands are shown in stick mode, coloured blue. The locations of the ligand-binding sites are highlighted with a yellow circle. Domain definitions were obtained from the Pfam database: 4HB, 4-helix bundle; sCache, single Cache; HBD, helical bimodular; dCache, double Cache. The dCache membrane-proximal and membrane-distal subdomains are labelled. PDB codes are shown in parentheses. Structure figures were prepared using PYMOL.

1.4.3.3. *H. pylori* chemoreceptors

For *H. pylori*, the ability to sense the external environment through chemical gradients is essential for its survival in the host environment and efficient colonisation of the gastric tissue. Previous studies have shown that *H. pylori* exhibits chemotaxis towards urea [213], bicarbonate [28], mucin, cholesterol [26], ZnCl2 [27], and several amino acids (alanine, arginine, asparagine, glutamine, histidine, proline, tyrosine, valine, leucine, serine, and glycine) [25,28]. In addition, *H. pylori* is repelled by aspartate [25], glutamate [25], AI-2 [212], $NiCl_2$ [27], low pH [210], and several bile acids (such as glycocholic, taurocholic, glycodeoxycholic, taurodeoxycholic, glycochenodeoxycholic, and taurochenodeoxycholic acid) [25].

In the *H. pylori* genome, four chemoreceptors or Tlps have been identified: TlpA, TlpB, TlpC, and TlpD (Figure 1.3, Table 1.1). TlpD is a soluble, cytoplasmic chemoreceptor composed of an MCP-signalling domain and a chemoreceptor zinc-binding (CZB) domain [232]. Although several cytoplasmic chemoreceptors harbour a CZB domain [232,233], the exact function of this domain in TlpD and other chemoreceptors remains unknown. TlpD is involved in energy taxis

[234] and mediates repellent responses away from conditions that promote oxidative stress, such as the presence of iron, hydrogen peroxide, paraquat, and metronidazole [235].

TlpA, TlpB, and TlpC are membrane-embedded proteins [30,211] (Figure 1.3). TlpA is a chemoreceptor that possesses a dCache LBD. It has been linked to the recognition of arginine and bicarbonate as attractants [28] and acidic pH as a repellent [236]. However, the mechanism by which these chemoeffectors stimulate the chemoreceptor is unknown. TlpB reportedly detects acidic pH and the quorum-sensing molecule AI-2 as repellents [210-212] and urea as an attractant [210,213]. TlpB directly binds urea with high affinity and specificity [210], and the crystal structure of its LBD in complex with this attractant has been solved. The TlpB LBD 3D structure revealed a homodimer of sCache modules, with a molecule of urea bound at the putative ligand-binding pocket of each sCache monomer [205,210]. Sensing of AI-2 by TlpB requires two periplasmic ligand-binding proteins, and therefore an indirect binding mechanism for detecting this signal has been suggested [214].

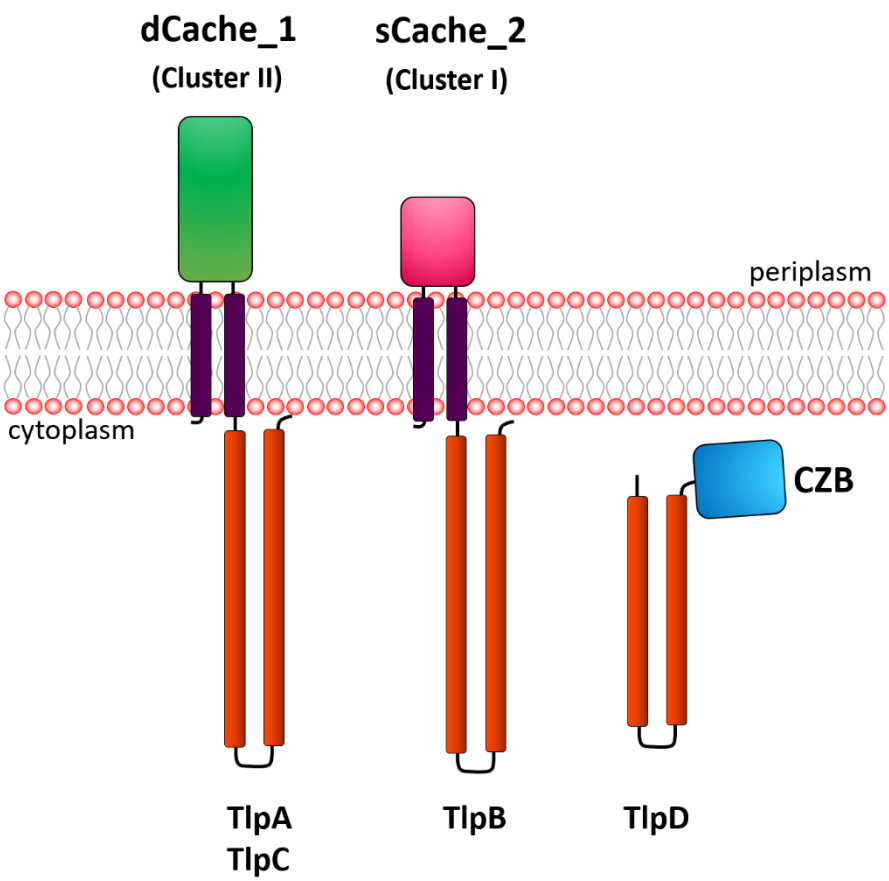


Figure 1.3. Set of chemoreceptors expressed by *H. pylori*. Receptor topology is shown for the four *H. pylori* Tlps. The LBD type is annotated for each chemoreceptor according to data from the Pfam database (http://pfam.xfam.org/).

Similarly to TlpA, TlpC also has a dCache LBD. Before the present research project was undertaken, TlpC was the least-characterised chemoreceptor in *H. pylori*. Furthermore, its ligand specificity had not been studied. Previous studies showed that mutation of *tlpC* reduced stomach colonisation [30] and that this chemoreceptor can modulate the TlpB-mediated response to acidic conditions [27].

1.4.3.4. *C. jejuni* chemoreceptors

In *C. jejuni*, chemotaxis is known to be critical for infection and colonisation of the host. The chemotactic behaviour of *C. jejuni* has been previously determined in the presence of different compounds [128]. Attractant responses were induced by several amino acids (such as aspartate, asparagine, cysteine, glutamate, isoleucine, and serine), carbohydrates (fucose and galactose), organic acids (lactic, malic, fumaric, formic, pyruvic, and succinic acid), sodium deoxycholate, and purine [128,237,238]. In contrast, *C. jejuni* exhibits a repellent response to lysine, arginine, glucosamine, thiamine, and bile constituents (cholic, doxycholic, glycocholic, glycochenodeoxycholic, glycodeoxycholic, and taurocholic acids) [128,238].

Analysis of the *C. jejuni* genomes of various laboratory and clinical strains has identified at least 15 different Tlps or chemoreceptors [179,239-242] (Figure 1.4, Table 1.1). These chemoreceptors have been categorised into three groups, A, B, and C, based on their domain organisation [179]. Group A receptors are membrane-embedded proteins consisting of an N-terminal transmembrane helix followed by a periplasmic LBD, a second transmembrane helix, and a C-terminal cytoplasmic signalling domain. Members of group A are Tlp1, Tlp2, Tlp3, Tlp4, Tlp7$_{mc}$, Tlp7$_{m}$, Tlp10, Tlp11, Tlp12, and Tlp13. Group B receptor Tlp9 is anchored to the inner membrane by two transmembrane helices, harbours cytoplasmic HAMP and signalling domains, but lacks a LBD. Finally, group C members, Tlp5, Tlp6, Tlp7$_{c}$ and Tlp8, are soluble cytoplasmic proteins that contain a signalling domain homologous to that in group A Tlps.

Table 1.1. Chemoreceptor repertoires of *H. pylori* and *C. jejuni.*

Chemoreceptor	LBD type[a]	Chemoeffectors	Comments
TlpA	dCache	Arginine, bicarbonate	Unknown binding mechanism [28]
TlpB	sCache	pH sensor	Unknown binding mechanism [210]
		AI-2	Indirect recognition [212]
		Urea, hydroxyurea, formamide, and acetamide	Direct recognition [210]
TlpC	dCache	Unknown	
TlpD	CZB	Energy [234] iron, hydrogen peroxide, paraquat, and metronidazole [235]	Unknown binding mechanism
Tlp1	dCache	Aspartate	Direct recognition [181]
Tlp2	dCache	Unknown	
Tlp3	dCache	Isoleucine, purine, aspartic acid, malic acid, fumaric acid, lysine, glucosamine, succinic acid, arginine, and thiamine	Direct recognition [237]
		Sodium deoxycholate	Unknown binding mechanism [238]
Tlp4	dCache	Sodium deoxycholate	Unknown binding mechanism [238]
Tlp5	Unidentifiable	Unknown	
Tlp6	CZB	Unknown	
Tlp7	Unidentifiable	Formic acid	Unknown binding mechanism [245]
Tlp8 (CetZ)	Tandem PAS_9	Energy	Unknown binding mechanism [244]
Tlp9 (CetA)	Lacks LBD	Energy	Unknown binding mechanism [244]
Tlp10	Unidentifiable	Unknown	
Tlp11	dCache	Galactose	Direct recognition [129,243]
Tlp12	Unidentifiable	Unknown	
Tlp13	Unidentifiable	Unknown	
CetB	PAS_3	Energy	Unknown binding mechanism [244]

[a]Domain predictions were made using Pfam and hmmsearch (https://www.ebi.ac.uk/Tools/hmmer/search/hmmsearch)

The ligand specificity has been established for only some of the *C. jejuni* chemoreceptors. From group A Tlps, five contain an identifiable dCache_1 LBD; these include Tlp1, Tlp2, Tlp3, Tlp4 and Tlp11 [193]. Tlp1 has been shown to recognise aspartate directly and was renamed the *Campylobacter* chemoreceptor for aspartate A (CcaA) [181]. Tlp3 has been shown to bind both attractants (isoleucine, purine, aspartic acid, malic acid, fumaric acid, and sodium deoxycholate) and repellents (lysine, glucosamine, succinic acid, arginine, and thiamine) [237,238]. Tlp4 appears to mediate chemotaxis to deoxycholate, the major component of bile [237], and Tlp11 has recently been identified as a galactose receptor [129,243].

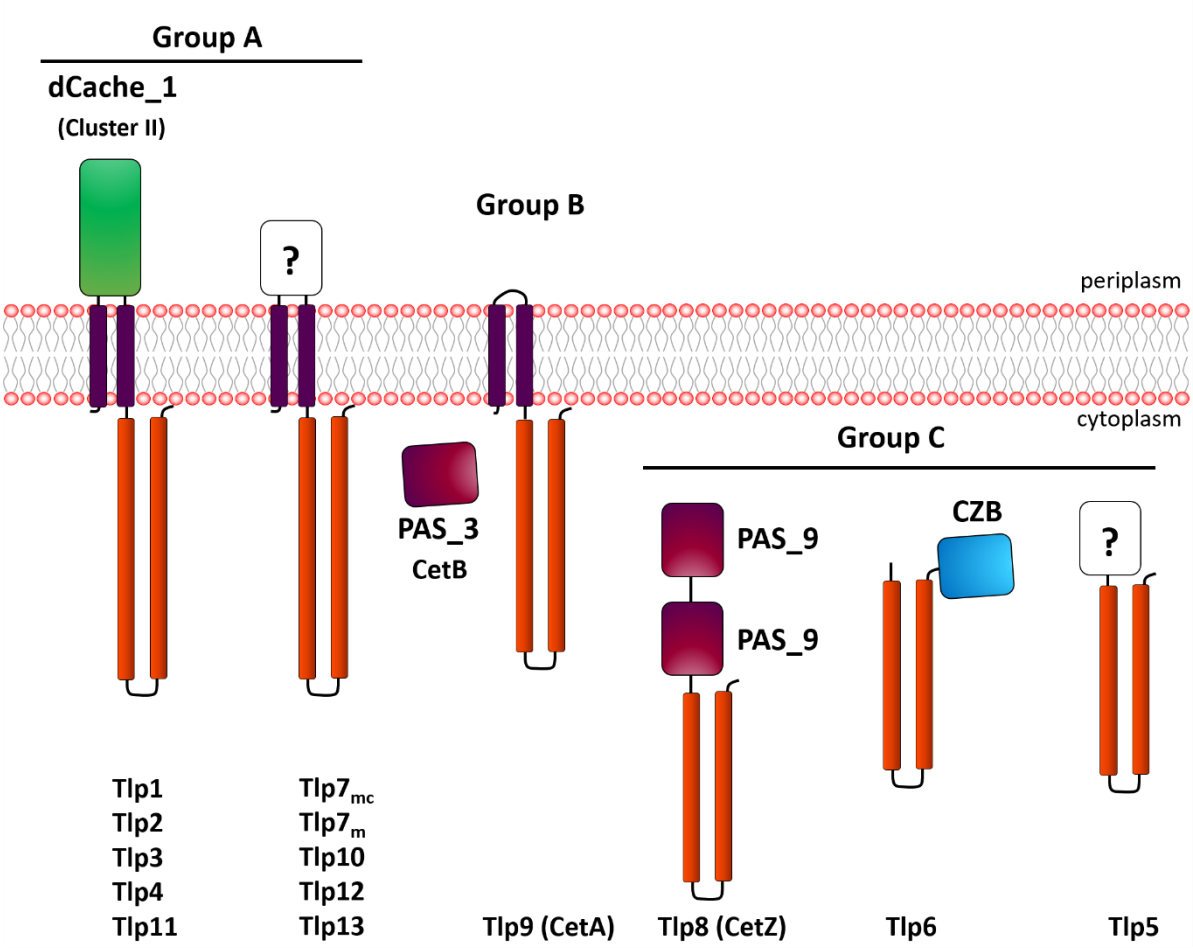


Figure 1.4. Chemoreceptor repertoire of *C. jejuni*. Receptor topology is shown for group A, B, and C chemoreceptors of *C. jejuni.* The identified LBDs are labelled according to the Pfam tool HMMER (http://pfam.xfam.org/).

Interestingly, *C. jejuni* has a bipartite chemoreceptor encoded by separate genes, *cetA* and *cetB* (*cet: **C**ampylobacter* **e**nergy **t**axis) [132,179]. The proteins produced by these genes interact in tandem, working as a fully functional chemoreceptor. *cetA* encodes the Tlp9

receptor, also known as CetA, which lacks a LBD, whereas *cetB* encodes a cytoplasmic PAS-containing protein termed CetB [179]. CetA/CetB is an energy taxis system which directs *C. jejuni* towards favourable conditions for energy generation and high redox potential. Tlp8, also known as CetZ, is a cytoplasmic receptor with two tandem PAS domains. In contrast to CetA/CetB, CetZ drives *C. jejuni* away from high redox potential [244].

Tlp7 is a receptor that mediates the attractant response to formic acid [245]. In most of *C. jejuni* isolates, the *tlp7* gene encodes a single transmembrane chemoreceptor of 58 kDa ($Tlp7_{mc}$). However, in some *C. jejuni* isolates, *tlp7* is interrupted by a stop codon that splits the gene into two ORFs, *cj0952c* and *cj0951c*. These two genes are expressed as independent proteins termed $Tlp7_c$ (Cj0951c) and $Tlp7_m$ (Cj0952c). $Tlp7_m$ is a membrane-anchored protein with a periplasmic LBD; in contrast, Tlp7c is a soluble protein with a cytoplasm signalling domain. When both proteins are expressed, they interact with one another as a heterodimer and mediate chemotaxis to formic acid, similarly to the full-length protein Tlp7 [245,246].

1.5. Project Aims

As reviewed above, chemotaxis and motility have been identified as important virulence factors for both *C. jejuni* and *H. pylori*, and are required for both colonisation and infection of their hosts [32,132,247]. Chemotaxis is the basis of directed movement in response to environmental cues, which allows motile bacteria to find optimum conditions for their proliferation. The sensors in bacterial chemotaxis are a diverse group of chemoreceptors, also known as methyl-accepting chemotaxis proteins (MCPs) or transducer-like proteins (Tlps), which recognise signal molecules through their ligand binding domains (LBDs) and transmit the information to cytoplasmic components [240].

Much of what is known about bacterial chemoreceptors originated from studies of *E. coli* Tsr and Tar receptors, which belong to subclass Ia (cluster I) and harbour a 4-helix bundle (4HB) LBD [197,248]. However, the LBDs of chemoreceptors are diverse in terms of their amino acid sequences and three-dimensional structures; to date, several structural families have been identified, including the helical biomodular (HBM), Per-ARNT-Sim (PAS), and the most abundant LBD type, the calcium channel and chemotaxis (Cache) superfamily.

The roles of some *C. jejuni* and *H. pylori* chemoreceptors in chemotaxis have been elucidated. For instance, of the fifteen chemoreceptors from *C. jejuni* and four from *H. pylori*,

only Tlp1 [181], Tlp3 [238], Tlp4 [237], Tlp7 [245], Tlp8 [244], Tlp9 [244], TlpA [249], TlpB [212], and TlpD [234] have been characterised and their ligand specificity identified. However, the molecular mechanism by which Tlps from these pathogens recognise environmental cues and discriminate between chemoattractants and repellents remains unclear.

In order to further understand the chemotaxis process in these closely related pathogens, this research aims to structurally characterise the LBDs of three chemoreceptors: Tlp1 and Tlp3 from *C. jejuni* and TlpC from *H. pylori.* Tlp1 and Tlp3 are the two most-characterised chemoreceptors from *C. jejuni*, and their function in chemotaxis has been investigated in depth. Contrastingly, little is known regarding the TlpC receptor, and its ligand specificity remains to be established. A comprehensive structural analysis of the LBDs of these chemoreceptors will provide molecular insights into the ligand recognition mechanisms employ by the Tlps, which may be exploited for the development of novel therapeutics and applications in industry (bioremediation).

Therefore the aims of this study are as follows:

Aim 1. Determination and analysis of the crystal structure of the dCache LBD of Tlp1 and Tlp3 from *C. jejuni* and TlpC from *H. pylori* in their free and chemoeffector-bound forms.

Aim 2. Determine the role of specific amino acids in Tlp1, Tlp3, and TlpC involved in chemoeffector binding affinity and specificity using site-directed mutagenesis and isothermal titration calorimetry.

Aim 3. To identify novel ligands forTlp3 using thermal shift assays and to analyse the crystal structure of Tlp3 in complex with such ligands.

1.6. References

1 Marshall, B. J. & Warren, J. R. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* **1**, 1311-1315 (1984).

2 Varbanova, M., Schulz, C. & Malfertheiner, P. Helicobacter pylori and other gastric bacteria. *Dig Dis* **29**, 562-569, doi:10.1159/000332989 (2011).

3 Suerbaum, S. & Michetti, P. Helicobacter pylori infection. *N Engl J Med* **347**, 1175-1186, doi:10.1056/NEJMra020542 (2002).

4 Kim, N. in *Helicobacter pylori* (ed N. Kim) Ch. 1, 3-19 (Springer, 2016).

5 Hooi, J. K. Y. *et al.* Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. *Gastroenterology* **153**, 420-429, doi:10.1053/j.gastro.2017.04.022 (2017).

6 Raymond, J. *et al.* Genetic and transmission analysis of Helicobacter pylori strains within a family. *Emerg Infect Dis* **10**, 1816-1821, doi:10.3201/eid1010.040042 (2004).

7 Didelot, X. *et al.* Genomic evolution and transmission of Helicobacter pylori in two South African families. *Proc Natl Acad Sci U S A* **110**, 13880-13885, doi:10.1073/pnas.1304681110 (2013).

8 Porsch-Ozcurumez, M. *et al.* Impact of migration on Helicobacter pylori seroprevalence in the offspring of Turkish immigrants in Germany. *Turk J Pediatr* **45**, 203-208 (2003).

9 Mitchell, H. M. in *Helicobacter pylori: Physiology and Genetics* (eds H. L. T. Mobley, G. L. Mendz, & S. L. Hazell) (2001).

10 Cover, T. L. & Blaser, M. J. Helicobacter pylori infection, a paradigm for chronic mucosal inflammation: pathogenesis and implications for eradication and prevention. *Adv Intern Med* **41**, 85-117 (1996).

11 Correa, P. & Piazuelo, M. B. Helicobacter pylori Infection and Gastric Adenocarcinoma. *US Gastroenterol Hepatol Rev* **7**, 59-64 (2011).

12 Malfertheiner, P., Chan, F. K. & McColl, K. E. Peptic ulcer disease. *Lancet* **374**, 1449-1461, doi:10.1016/S0140-6736(09)60938-7 (2009).

13 Peek, R. M., Jr. & Blaser, M. J. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nat. Rev. Cancer* **2**, 28-37, doi:10.1038/nrc703 (2002).

14 Parkin, D. M., Bray, F., Ferlay, J. & Pisani, P. Global cancer statistics, 2002. *CA Cancer J Clin* **55**, 74-108 (2005).

15 Plummer, M., Franceschi, S., Vignat, J., Forman, D. & de Martel, C. Global burden of gastric cancer attributable to Helicobacter pylori. *Int J Cancer* **136**, 487-490, doi:10.1002/ijc.28999 (2015).

16 Frydman, G. H., Davis, N., Beck, P. L. & Fox, J. G. Helicobacter pylori Eradication in Patients with Immune Thrombocytopenic Purpura: A Review and the Role of Biogeography. *Helicobacter* **20**, 239-251, doi:10.1111/hel.12200 (2015).

17 Chiba, N. *et al.* Treating Helicobacter pylori infection in primary care patients with uninvestigated dyspepsia: the Canadian adult dyspepsia empiric treatment-Helicobacter pylori positive (CADET-Hp) randomised controlled trial. *BMJ* **324**, 1012-1016 (2002).

18 O'Toole, P. W., Lane, M. C. & Porwollik, S. Helicobacter pylori motility. *Microbes Infect* **2**, 1207-1214 (2000).

19 Geis, G., Leying, H., Suerbaum, S., Mai, U. & Opferkuch, W. Ultrastructure and chemical analysis of Campylobacter pylori flagella. *J Clin Microbiol* **27**, 436-441 (1989).

20 Eaton, K. A., Morgan, D. R. & Krakowka, S. Motility as a factor in the colonisation of gnotobiotic piglets by Helicobacter pylori. *J Med Microbiol* **37**, 123-127, doi:10.1099/00222615-37-2-123 (1992).

21 Eaton, K. A., Suerbaum, S., Josenhans, C. & Krakowka, S. Colonization of gnotobiotic piglets by Helicobacter pylori deficient in two flagellin genes. *Infect Immun* **64**, 2445-2448 (1996).

22 Foynes, S. *et al.* Functional analysis of the roles of FliQ and FlhB in flagellar expression in Helicobacter pylori. *FEMS Microbiol Lett* **174**, 33-39 (1999).

23 Ottemann, K. M. & Lowenthal, A. C. *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. *Infection and immunity* **70**, 1984-1990 (2002).

24 Dunne, C., Dolan, B. & Clyne, M. Factors that mediate colonization of the human stomach by Helicobacter pylori. *World J Gastroenterol* **20**, 5610-5624, doi:10.3748/wjg.v20.i19.5610 (2014).

25 Worku, M. L., Karim, Q. N., Spencer, J. & Sidebotham, R. L. Chemotactic response of *Helicobacter pylori* to human plasma and bile. *J. Med. Microbiol.* **53**, 807-811, doi:10.1099/jmm.0.45636-0 (2004).

26 Wunder, C. *et al.* Cholesterol glucosylation promotes immune evasion by *Helicobacter pylori*. *Nature medicine* **12**, 1030-1038, doi:10.1038/nm1480 (2006).

27 Sanders, L., Andermann, T. M. & Ottemann, K. M. A supplemented soft agar chemotaxis assay demonstrates the *Helicobacter pylori* chemotactic response to zinc and nickel. *Microbiology* **159**, 46-57, doi:10.1099/mic.0.062877-0 (2013).

28 Cerda, O., Rivas, A. & Toledo, H. *Helicobacter pylori* strain ATCC700392 encodes a methyl-accepting chemotaxis receptor protein (MCP) for arginine and sodium bicarbonate. *FEMS microbiology letters* **224**, 175-181 (2003).

29 Machuca, M. A. *et al.* Helicobacter pylori chemoreceptor TlpC mediates chemotaxis to lactate. *Sci Rep* **7**, 14089, doi:10.1038/s41598-017-14372-2 (2017).

30 Andermann, T. M., Chen, Y. T. & Ottemann, K. M. Two predicted chemoreceptors of *Helicobacter pylori* promote stomach infection. *Infection and immunity* **70**, 5877-5881 (2002).

31 Foynes, S. *et al. Helicobacter pylori* possesses two CheY response regulators and a histidine kinase sensor, CheA, which are essential for chemotaxis and colonization of the gastric mucosa. *Infection and immunity* **68**, 2016-2023 (2000).

32 Terry, K., Williams, S. M., Connolly, L. & Ottemann, K. M. Chemotaxis plays multiple roles during *Helicobacter pylori* animal infection. *Infection and immunity* **73**, 803-811, doi:10.1128/IAI.73.2.803-811.2005 (2005).

33 Howitt, M. R. *et al.* ChePep controls Helicobacter pylori Infection of the gastric glands and chemotaxis in the Epsilonproteobacteria. *MBio* **2**, doi:10.1128/mBio.00098-11 (2011).

34 Williams, S. M. *et al.* Helicobacter pylori chemotaxis modulates inflammation and bacterium-gastric epithelium interactions in infected mice. *Infect Immun* **75**, 3747-3757, doi:10.1128/IAI.00082-07 (2007).

35 Eaton, K. A., Brooks, C. L., Morgan, D. R. & Krakowka, S. Essential role of urease in pathogenesis of gastritis induced by Helicobacter pylori in gnotobiotic piglets. *Infect Immun* **59**, 2470-2475 (1991).

36 Mobley, H. L., Island, M. D. & Hausinger, R. P. Molecular biology of microbial ureases. *Microbiol Rev* **59**, 451-480 (1995).

37 Clyne, M., Dolan, B. & Reeves, E. P. Bacterial factors that mediate colonization of the stomach and virulence of Helicobacter pylori. *FEMS Microbiol Lett* **268**, 135-143, doi:10.1111/j.1574-6968.2007.00648.x (2007).

38 Labigne, A., Cussac, V. & Courcoux, P. Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsible for urease activity. *J Bacteriol* **173**, 1920-1931 (1991).

39 Akada, J. K., Shirai, M., Takeuchi, H., Tsuda, M. & Nakazawa, T. Identification of the urease operon in Helicobacter pylori and its control by mRNA decay in response to pH. *Mol Microbiol* **36**, 1071-1084 (2000).

40 Fong, Y. H. *et al.* Structure of UreG/UreF/UreH complex reveals how urease accessory proteins facilitate maturation of Helicobacter pylori urease. *PLoS Biol* **11**, e1001678, doi:10.1371/journal.pbio.1001678 (2013).

41 Ha, N. C. *et al.* Supramolecular assembly and acid resistance of Helicobacter pylori urease. *Nat Struct Biol* **8**, 505-509, doi:10.1038/88563 (2001).

42 de Reuse, H., Vinella, D. & Cavazza, C. Common themes and unique proteins for the uptake and trafficking of nickel, a metal essential for the virulence of Helicobacter pylori. *Front Cell Infect Microbiol* **3**, 94, doi:10.3389/fcimb.2013.00094 (2013).

43 Blusiewicz, K., Rydzewska, G. & Rydzewski, A. Gastric juice ammonia and urea concentrations and their relation to gastric mucosa injury in patients maintained on chronic hemodialysis. *Rocz Akad Med Bialymst* **50**, 188-192 (2005).

44 Weeks, D. L., Eskandari, S., Scott, D. R. & Sachs, G. A H+-gated urea channel: the link between Helicobacter pylori urease and gastric colonization. *Science* **287**, 482-485 (2000).

45 Gray, L. R., Gu, S. X., Quick, M. & Khademi, S. Transport kinetics and selectivity of HpUreI, the urea channel from Helicobacter pylori. *Biochemistry* **50**, 8656-8663, doi:10.1021/bi200887a (2011).

46 Scott, D. R. *et al.* Expression of the Helicobacter pylori ureI gene is required for acidic pH activation of cytoplasmic urease. *Infect Immun* **68**, 470-477 (2000).

47 Bury-Mone, S., Skouloubris, S., Labigne, A. & De Reuse, H. The Helicobacter pylori UreI protein: role in adaptation to acidity and identification of residues essential for its activity and for acid activation. *Mol Microbiol* **42**, 1021-1034 (2001).

48 Mollenhauer-Rektorschek, M., Hanauer, G., Sachs, G. & Melchers, K. Expression of UreI is required for intragastric transit and colonization of gerbil gastric mucosa by Helicobacter pylori. *Res Microbiol* **153**, 659-666 (2002).

49 Skouloubris, S., Thiberge, J. M., Labigne, A. & De Reuse, H. The Helicobacter pylori UreI protein is not involved in urease activity but is essential for bacterial survival in vivo. *Infect Immun* **66**, 4517-4521 (1998).

50 Ilver, D. *et al.* Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* **279**, 373-377 (1998).

51 Boren, T., Falk, P., Roth, K. A., Larson, G. & Normark, S. Attachment of Helicobacter pylori to human gastric epithelium mediated by blood group antigens. *Science* **262**, 1892-1895 (1993).

52 Ishijima, N. *et al.* BabA-mediated adherence is a potentiator of the Helicobacter pylori type IV secretion system activity. *J Biol Chem* **286**, 25256-25264, doi:10.1074/jbc.M111.233601 (2011).

53 Mahdavi, J. *et al.* Helicobacter pylori SabA adhesin in persistent infection and chronic inflammation. *Science* **297**, 573-578, doi:10.1126/science.1069076 (2002).

54 Yamaoka, Y. Increasing evidence of the role of Helicobacter pylori SabA in the pathogenesis of gastroduodenal disease. *J Infect Dev Ctries* **2**, 174-181 (2008).

55 Odenbreit, S., Till, M., Hofreuter, D., Faller, G. & Haas, R. Genetic and functional characterization of the alpAB gene locus essential for the adhesion of Helicobacter pylori to human gastric tissue. *Mol Microbiol* **31**, 1537-1548 (1999).

56 Odenbreit, S., Faller, G. & Haas, R. Role of the alpAB proteins and lipopolysaccharide in adhesion of Helicobacter pylori to human gastric tissue. *Int J Med Microbiol* **292**, 247-256, doi:10.1078/1438-4221-00204 (2002).

57 Dossumbekova, A. *et al.* Helicobacter pylori HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of hopH gene polymorphisms. *J Infect Dis* **194**, 1346-1355, doi:10.1086/508426 (2006).

58 Yamaoka, Y. *et al.* Importance of Helicobacter pylori oipA in clinical presentation, gastric inflammation, and mucosal interleukin 8 production. *Gastroenterology* **123**, 414-424 (2002).

59 Backert, S., Clyne, M. & Tegtmeyer, N. Molecular mechanisms of gastric epithelial cell adhesion and injection of CagA by Helicobacter pylori. *Cell Commun Signal* **9**, 28, doi:10.1186/1478-811X-9-28 (2011).

60 Franco, A. T. *et al.* Regulation of gastric carcinogenesis by Helicobacter pylori virulence factors. *Cancer Res* **68**, 379-387, doi:10.1158/0008-5472.CAN-07-0824 (2008).

61 Raetz, C. R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**, 635-700, doi:10.1146/annurev.biochem.71.110601.135414 (2002).

62 Moran, A. P., Lindner, B. & Walsh, E. J. Structural characterization of the lipid A component of Helicobacter pylori rough- and smooth-form lipopolysaccharides. *J Bacteriol* **179**, 6453-6463 (1997).

63 Aspinall, G. O., Monteiro, M. A., Pang, H., Walsh, E. J. & Moran, A. P. Lipopolysaccharide of the Helicobacter pylori type strain NCTC 11637 (ATCC 43504): structure of the O antigen chain and core oligosaccharide regions. *Biochemistry* **35**, 2489-2497, doi:10.1021/bi951852s (1996).

64 Moran, A. P. Biological and serological characterization of Campylobacter jejuni lipopolysaccharides with deviating core and lipid A structures. *FEMS Immunol Med Microbiol* **11**, 121-130 (1995).

65 Muotiala, A., Helander, I. M., Pyhala, L., Kosunen, T. U. & Moran, A. P. Low biological activity of Helicobacter pylori lipopolysaccharide. *Infect Immun* **60**, 1714-1716 (1992).

66 Moran, A. P. Lipopolysaccharide in bacterial chronic infection: insights from Helicobacter pylori lipopolysaccharide and lipid A. *Int J Med Microbiol* **297**, 307-319, doi:10.1016/j.ijmm.2007.03.008 (2007).

67 Simoons-Smit, I. M. *et al.* Typing of Helicobacter pylori with monoclonal antibodies against Lewis antigens in lipopolysaccharide. *J Clin Microbiol* **34**, 2196-2200 (1996).

68 Heneghan, M. A., McCarthy, C. F. & Moran, A. P. Relationship of blood group determinants on Helicobacter pylori lipopolysaccharide with host lewis phenotype and inflammatory response. *Infect Immun* **68**, 937-941 (2000).

69 Monteiro, M. A. *et al.* Simultaneous expression of type 1 and type 2 Lewis blood group antigens by Helicobacter pylori lipopolysaccharides. Molecular mimicry between h.

pylori lipopolysaccharides and human gastric epithelial cell surface glycoforms. *J Biol Chem* **273**, 11533-11543 (1998).

70 Moran, A. P. in *Helicobacter pylori: Physiology and Genetics* (eds H. L. T. Mobley, G. L. Mendz, & S. L. Hazell) (2001).

71 Appelmelk, B. J., Monteiro, M. A., Martin, S. L., Moran, A. P. & Vandenbroucke-Grauls, C. M. Why Helicobacter pylori has Lewis antigens. *Trends Microbiol* **8**, 565-570 (2000).

72 Wirth, H. P., Yang, M., Peek, R. M., Jr., Tham, K. T. & Blaser, M. J. Helicobacter pylori Lewis expression is related to the host Lewis phenotype. *Gastroenterology* **113**, 1091-1098 (1997).

73 Hynes, S. O., Keenan, J. I., Ferris, J. A., Annuk, H. & Moran, A. P. Lewis epitopes on outer membrane vesicles of relevance to Helicobacter pylori pathogenesis. *Helicobacter* **10**, 146-156, doi:10.1111/j.1523-5378.2005.00302.x (2005).

74 Appelmelk, B. J. *et al.* Potential role of molecular mimicry between Helicobacter pylori lipopolysaccharide and host Lewis blood group antigens in autoimmunity. *Infect Immun* **64**, 2031-2040 (1996).

75 Heneghan, M. A., McCarthy, C. F., Janulaityte, D. & Moran, A. P. Relationship of anti-Lewis x and anti-Lewis y antibodies in serum samples from gastric cancer and chronic gastritis patients to Helicobacter pylori-mediated autoimmunity. *Infect Immun* **69**, 4774-4781, doi:10.1128/IAI.69.8.4774-4781.2001 (2001).

76 Fowler, M., Thomas, R. J., Atherton, J., Roberts, I. S. & High, N. J. Galectin-3 binds to Helicobacter pylori O-antigen: it is upregulated and rapidly secreted by gastric epithelial cells in response to H. pylori adhesion. *Cell Microbiol* **8**, 44-54, doi:10.1111/j.1462-5822.2005.00599.x (2006).

77 Edwards, N. J. *et al.* Lewis X structures in the O antigen side-chain promote adhesion of Helicobacter pylori to the gastric epithelium. *Mol Microbiol* **35**, 1530-1539 (2000).

78 Montecucco, C. & Rappuoli, R. Living dangerously: how Helicobacter pylori survives in the human stomach. *Nature reviews. Molecular cell biology* **2**, 457-466, doi:10.1038/35073084 (2001).

79 Yahiro, K. *et al.* Activation of Helicobacter pylori VacA toxin by alkaline or acid conditions increases its binding to a 250-kDa receptor protein-tyrosine phosphatase beta. *J Biol Chem* **274**, 36693-36699 (1999).

80 Yahiro, K. *et al.* Protein-tyrosine phosphatase alpha, RPTP alpha, is a Helicobacter pylori VacA receptor. *J Biol Chem* **278**, 19183-19189, doi:10.1074/jbc.M300117200 (2003).

81 Gupta, V. R. *et al.* Sphingomyelin functions as a novel receptor for Helicobacter pylori VacA. *PLoS Pathog* **4**, e1000073, doi:10.1371/journal.ppat.1000073 (2008).

82 Seto, K., Hayashi-Kuwabara, Y., Yoneta, T., Suda, H. & Tamaki, H. Vacuolation induced by cytotoxin from Helicobacter pylori is mediated by the EGF receptor in HeLa cells. *FEBS Lett* **431**, 347-350 (1998).

83 Utt, M., Danielsson, B. & Wadstrom, T. Helicobacter pylori vacuolating cytotoxin binding to a putative cell surface receptor, heparan sulfate, studied by surface plasmon resonance. *FEMS Immunol Med Microbiol* **30**, 109-113 (2001).

84 Yahiro, K. *et al.* Low-density lipoprotein receptor-related protein-1 (LRP1) mediates autophagy and apoptosis caused by Helicobacter pylori VacA. *J Biol Chem* **287**, 31104-31115, doi:10.1074/jbc.M112.387498 (2012).

85 Roche, N. *et al.* Human gastric glycosphingolipids recognized by Helicobacter pylori vacuolating cytotoxin VacA. *Microbes Infect* **9**, 605-614, doi:10.1016/j.micinf.2007.01.023 (2007).

86 Ricci, V. *et al.* High cell sensitivity to Helicobacter pylori VacA toxin depends on a GPI-anchored protein and is not blocked by inhibition of the clathrin-mediated pathway of endocytosis. *Mol Biol Cell* **11**, 3897-3909 (2000).

87 Ricci, V. *et al.* Helicobacter pylori vacuolating toxin accumulates within the endosomal-vacuolar compartment of cultured gastric cells and potentiates the vacuolating activity of ammonia. *J Pathol* **183**, 453-459, doi:10.1002/(SICI)1096-9896(199712)183:4<453::AID-PATH950>3.0.CO;2-2 (1997).

88 Papini, E. *et al.* The vacuolar ATPase proton pump is present on intracellular vacuoles induced by Helicobacter pylori. *J Med Microbiol* **45**, 84-89, doi:10.1099/00222615-45-2-84 (1996).

89 Molinari, M. *et al.* Selective inhibition of Ii-dependent antigen presentation by Helicobacter pylori toxin VacA. *J Exp Med* **187**, 135-140 (1998).

90 Satin, B. *et al.* Effect of helicobacter pylori vacuolating toxin on maturation and extracellular release of procathepsin D and on epidermal growth factor degradation. *J Biol Chem* **272**, 25022-25028 (1997).

91 Testerman, T. L. & Morris, J. Beyond the stomach: an updated view of Helicobacter pylori pathogenesis, diagnosis, and treatment. *World J Gastroenterol* **20**, 12781-12808, doi:10.3748/wjg.v20.i36.12781 (2014).

92 Papini, E. *et al.* Selective increase of the permeability of polarized epithelial cell monolayers by Helicobacter pylori vacuolating toxin. *J Clin Invest* **102**, 813-820, doi:10.1172/JCI2764 (1998).

93 Tombola, F. *et al.* The Helicobacter pylori VacA toxin is a urea permease that promotes urea diffusion across epithelia. *J Clin Invest* **108**, 929-937, doi:10.1172/JCI13045 (2001).

94 Lu, H., Yamaoka, Y. & Graham, D. Y. Helicobacter pylori virulence factors: facts and fantasies. *Current opinion in gastroenterology* **21**, 653-659 (2005).

95 Censini, S. *et al.* cag, a pathogenicity island of Helicobacter pylori, encodes type I-specific and disease-associated virulence factors. *Proc Natl Acad Sci U S A* **93**, 14648-14653 (1996).

96 Noto, J. M. & Peek, R. M., Jr. The Helicobacter pylori cag Pathogenicity Island. *Methods Mol Biol* **921**, 41-50, doi:10.1007/978-1-62703-005-2_7 (2012).

97 Odenbreit, S. *et al.* Translocation of Helicobacter pylori CagA into gastric epithelial cells by type IV secretion. *Science* **287**, 1497-1500 (2000).

98 Amieva, M. R. *et al.* Disruption of the epithelial apical-junctional complex by Helicobacter pylori CagA. *Science* **300**, 1430-1434, doi:10.1126/science.1081919 (2003).

99 Hatakeyama, M. Oncogenic mechanisms of the Helicobacter pylori CagA protein. *Nature reviews. Cancer* **4**, 688-694, doi:10.1038/nrc1433 (2004).

100 Backert, S., Tegtmeyer, N. & Selbach, M. The versatility of Helicobacter pylori CagA effector protein functions: The master key hypothesis. *Helicobacter* **15**, 163-176, doi:10.1111/j.1523-5378.2010.00759.x (2010).

101 Murata-Kamiya, N. *et al.* Helicobacter pylori CagA interacts with E-cadherin and deregulates the beta-catenin signal that promotes intestinal transdifferentiation in gastric epithelial cells. *Oncogene* **26**, 4617-4626, doi:10.1038/sj.onc.1210251 (2007).

102 Nesic, D. *et al.* Helicobacter pylori CagA inhibits PAR1-MARK family kinases by mimicking host substrates. *Nat Struct Mol Biol* **17**, 130-132, doi:10.1038/nsmb.1705 (2010).

103 Ohnishi, N. *et al.* Transgenic expression of Helicobacter pylori CagA induces gastrointestinal and hematopoietic neoplasms in mouse. *Proc Natl Acad Sci U S A* **105**, 1003-1008, doi:10.1073/pnas.0711183105 (2008).

104 Snelling, W. J., Matsuda, M., Moore, J. E. & Dooley, J. S. Campylobacter jejuni. *Lett Appl Microbiol* **41**, 297-302, doi:10.1111/j.1472-765X.2005.01788.x (2005).

105 Boes, J. *et al.* Prevalence and diversity of Campylobacter jejuni in pig herds on farms with and without cattle or poultry. *J Food Prot* **68**, 722-727 (2005).

106 Oporto, B., Esteban, J. I., Aduriz, G., Juste, R. A. & Hurtado, A. Prevalence and strain diversity of thermophilic campylobacters in cattle, sheep and swine farms. *J Appl Microbiol* **103**, 977-984, doi:10.1111/j.1365-2672.2007.03328.x (2007).

107 Petersen, L., Nielsen, E. M., Engberg, J., On, S. L. & Dietz, H. H. Comparison of genotypes and serotypes of Campylobacter jejuni isolated from Danish wild mammals and birds and from broiler flocks and humans. *Appl Environ Microbiol* **67**, 3115-3121, doi:10.1128/AEM.67.7.3115-3121.2001 (2001).

108 Allos, B. M. Campylobacter jejuni Infections: update on emerging issues and trends. *Clin Infect Dis* **32**, 1201-1206, doi:10.1086/319760 (2001).

109 Shane, S. M. Campylobacter infection of commercial poultry. *Rev Sci Tech* **19**, 376-395 (2000).

110 van Vliet, A. H. & Ketley, J. M. Pathogenesis of enteric Campylobacter infection. *Symposium series*, 45S-56S (2001).

111 Butzler, J. P., Dekeyser, P., Detrain, M. & Dehaen, F. Related vibrio in stools. *J Pediatr* **82**, 493-495 (1973).

112 Skirrow, M. B. Campylobacter enteritis: a "new" disease. *Br Med J* **2**, 9-11 (1977).

113 Epps, S. V. *et al.* Foodborne Campylobacter: infections, metabolism, pathogenesis and reservoirs. *Int J Environ Res Public Health* **10**, 6292-6304, doi:10.3390/ijerph10126292 (2013).

114 Crim, S. M. *et al.* Incidence and trends of infection with pathogens transmitted commonly through food--Foodborne Diseases Active Surveillance Network, 10 U.S. sites, 2006-2013. *MMWR Morb Mortal Wkly Rep* **63**, 328-332 (2014).

115 Kaakoush, N. O., Castano-Rodriguez, N., Mitchell, H. M. & Man, S. M. Global Epidemiology of Campylobacter Infection. *Clin Microbiol Rev* **28**, 687-720, doi:10.1128/CMR.00006-15 (2015).

116 Arsenault, J., Berke, O., Michel, P., Ravel, A. & Gosselin, P. Environmental and demographic risk factors for campylobacteriosis: do various geographical scales tell the same story? *BMC Infect Dis* **12**, 318, doi:10.1186/1471-2334-12-318 (2012).

117 Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: annual report of the OzFoodNet network, 2010., (OzFoodNet Working Group, 2012).

118 Platts-Mills, J. A. *et al.* Detection of Campylobacter in stool and determination of significance by culture, enzyme immunoassay, and PCR in developing countries. *J Clin Microbiol* **52**, 1074-1080, doi:10.1128/JCM.02935-13 (2014).

119 Blaser, M. J. Epidemiologic and clinical features of Campylobacter jejuni infections. *J Infect Dis* **176 Suppl 2**, S103-105 (1997).

120 Kuroki, S. *et al.* Guillain-Barre syndrome associated with Campylobacter infection. *Pediatr Infect Dis J* **10**, 149-151 (1991).

121 Ang, C. W. *et al.* Guillain-Barre syndrome- and Miller Fisher syndrome-associated Campylobacter jejuni lipopolysaccharides induce anti-GM1 and anti-GQ1b Antibodies in rabbits. *Infect Immun* **69**, 2462-2469, doi:10.1128/IAI.69.4.2462-2469.2001 (2001).

122 Lee, A., O'Rourke, J. L., Barrington, P. J. & Trust, T. J. Mucus colonization as a determinant of pathogenicity in intestinal infection by Campylobacter jejuni: a mouse cecal model. *Infect Immun* **51**, 536-546 (1986).

123 Wassenaar, T. M., Bleumink-Pluym, N. M. & van der Zeijst, B. A. Inactivation of Campylobacter jejuni flagellin genes by homologous recombination demonstrates that flaA but not flaB is required for invasion. *EMBO J* **10**, 2055-2061 (1991).

124 Nachamkin, I., Yang, X. H. & Stern, N. J. Role of Campylobacter jejuni flagella as colonization factors for three-day-old chicks: analysis with flagellar mutants. *Appl Environ Microbiol* **59**, 1269-1273 (1993).

125 Black, R. E., Levine, M. M., Clements, M. L., Hughes, T. P. & Blaser, M. J. Experimental Campylobacter jejuni infection in humans. *J Infect Dis* **157**, 472-479 (1988).

126 Yao, R. *et al.* Isolation of motile and non-motile insertional mutants of Campylobacter jejuni: the role of motility in adherence and invasion of eukaryotic cells. *Mol Microbiol* **14**, 883-893 (1994).

127 Szymanski, C. M., King, M., Haardt, M. & Armstrong, G. D. Campylobacter jejuni motility and invasion of Caco-2 cells. *Infect Immun* **63**, 4295-4300 (1995).

128 Hugdahl, M. B., Beery, J. T. & Doyle, M. P. Chemotactic behavior of *Campylobacter jejuni*. *Infection and immunity* **56**, 1560-1566 (1988).

129 Day, C. J. *et al.* A direct-sensing galactose chemoreceptor recently evolved in invasive strains of Campylobacter jejuni. *Nat Commun* **7**, 13206, doi:10.1038/ncomms13206 (2016).

130 Vegge, C. S., Brondsted, L., Li, Y. P., Bang, D. D. & Ingmer, H. Energy taxis drives Campylobacter jejuni toward the most favorable conditions for growth. *Appl Environ Microbiol* **75**, 5308-5314, doi:10.1128/AEM.00287-09 (2009).

131 Morooka, T., Umeda, A. & Amako, K. Motility as an intestinal colonization factor for *Campylobacter jejuni*. *J Gen Microbiol* **131**, 1973-1980, doi:10.1099/00221287-131-8-1973 (1985).

132 Hendrixson, D. R. & DiRita, V. J. Identification of Campylobacter jejuni genes involved in commensal colonization of the chick gastrointestinal tract. *Mol Microbiol* **52**, 471-484, doi:10.1111/j.1365-2958.2004.03988.x (2004).

133 Chang, C. & Miller, J. F. Campylobacter jejuni colonization of mice with limited enteric flora. *Infect Immun* **74**, 5261-5271, doi:10.1128/IAI.01094-05 (2006).

134 Konkel, M. E., Kim, B. J., Rivera-Amill, V. & Garvis, S. G. Bacterial secreted proteins are required for the internaliztion of Campylobacter jejuni into cultured mammalian cells. *Mol Microbiol* **32**, 691-701 (1999).

135 Christensen, J. E., Pacheco, S. A. & Konkel, M. E. Identification of a Campylobacter jejuni-secreted protein required for maximal invasion of host cells. *Mol Microbiol* **73**, 650-662, doi:10.1111/j.1365-2958.2009.06797.x (2009).

136 Neal-McKinney, J. M., Christensen, J. E. & Konkel, M. E. Amino-terminal residues dictate the export efficiency of the Campylobacter jejuni filament proteins via the flagellum. *Mol Microbiol* **76**, 918-931, doi:10.1111/j.1365-2958.2010.07144.x (2010).

137 Konkel, M. E. *et al.* Secretion of virulence proteins from Campylobacter jejuni is dependent on a functional flagellar export apparatus. *J Bacteriol* **186**, 3296-3303, doi:10.1128/JB.186.11.3296-3303.2004 (2004).

138 Eucker, T. P. & Konkel, M. E. The cooperative action of bacterial fibronectin-binding proteins and secreted proteins promote maximal Campylobacter jejuni invasion of host cells by stimulating membrane ruffling. *Cell Microbiol* **14**, 226-238, doi:10.1111/j.1462-5822.2011.01714.x (2012).

139 Barrero-Tobon, A. M. & Hendrixson, D. R. Identification and analysis of flagellar coexpressed determinants (Feds) of Campylobacter jejuni involved in colonization. *Mol Microbiol* **84**, 352-369, doi:10.1111/j.1365-2958.2012.08027.x (2012).

140 Buelow, D. R., Christensen, J. E., Neal-McKinney, J. M. & Konkel, M. E. Campylobacter jejuni survival within human epithelial cells is enhanced by the secreted protein CiaI. *Mol Microbiol* **80**, 1296-1312, doi:10.1111/j.1365-2958.2011.07645.x (2011).

141 Poly, F. & Guerry, P. Pathogenesis of Campylobacter. *Current opinion in gastroenterology* **24**, 27-31, doi:10.1097/MOG.0b013e3282f1dcb1 (2008).

142 Karlyshev, A. V. *et al.* Analysis of Campylobacter jejuni capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol Microbiol* **55**, 90-103, doi:10.1111/j.1365-2958.2004.04374.x (2005).

143 Bacon, D. J. *et al.* A phase-variable capsule is involved in virulence of Campylobacter jejuni 81-176. *Mol Microbiol* **40**, 769-777 (2001).

144 Champion, O. L. *et al.* Insect infection model for Campylobacter jejuni reveals that O-methyl phosphoramidate has insecticidal activity. *J Infect Dis* **201**, 776-782, doi:10.1086/650494 (2010).

145 Rose, A., Kay, E., Wren, B. W. & Dallman, M. J. The Campylobacter jejuni NCTC11168 capsule prevents excessive cytokine production by dendritic cells. *Med Microbiol Immunol* **201**, 137-144, doi:10.1007/s00430-011-0214-1 (2012).

146 Karlyshev, A. V., Ketley, J. M. & Wren, B. W. The Campylobacter jejuni glycome. *FEMS Microbiol Rev* **29**, 377-390, doi:10.1016/j.femsre.2005.01.003 (2005).

147 St Michael, F. *et al.* The structures of the lipooligosaccharide and capsule polysaccharide of Campylobacter jejuni genome sequenced strain NCTC 11168. *Eur J Biochem* **269**, 5119-5136 (2002).

148 Guerry, P., Ewing, C. P., Hickey, T. E., Prendergast, M. M. & Moran, A. P. Sialylation of lipooligosaccharide cores affects immunogenicity and serum resistance of Campylobacter jejuni. *Infect Immun* **68**, 6656-6662 (2000).

149 Konkel, M. E. *et al.* Identification of a fibronectin-binding domain within the Campylobacter jejuni CadF protein. *Mol Microbiol* **57**, 1022-1035, doi:10.1111/j.1365-2958.2005.04744.x (2005).

150 Konkel, M. E., Larson, C. L. & Flanagan, R. C. Campylobacter jejuni FlpA binds fibronectin and is required for maximal host cell adherence. *J Bacteriol* **192**, 68-76, doi:10.1128/JB.00969-09 (2010).

151 Jin, S., Song, Y. C., Emili, A., Sherman, P. M. & Chan, V. L. JlpA of Campylobacter jejuni interacts with surface-exposed heat shock protein 90alpha and triggers signalling pathways leading to the activation of NF-kappaB and p38 MAP kinase in epithelial cells. *Cell Microbiol* **5**, 165-174 (2003).

152 Jin, S. *et al.* JlpA, a novel surface-exposed lipoprotein specific to Campylobacter jejuni, mediates adherence to host epithelial cells. *Mol Microbiol* **39**, 1225-1236 (2001).

153 Pei, Z. & Blaser, M. J. PEB1, the major cell-binding factor of Campylobacter jejuni, is a homolog of the binding component in gram-negative nutrient transport systems. *J Biol Chem* **268**, 18717-18725 (1993).

154 Monteville, M. R., Yoon, J. E. & Konkel, M. E. Maximal adherence and invasion of INT 407 cells by Campylobacter jejuni requires the CadF outer-membrane protein and microfilament reorganization. *Microbiology* **149**, 153-165, doi:10.1099/mic.0.25820-0 (2003).

155 Krause-Gruszczynska, M. *et al.* Role of the small Rho GTPases Rac1 and Cdc42 in host cell invasion of Campylobacter jejuni. *Cell Microbiol* **9**, 2431-2444, doi:10.1111/j.1462-5822.2007.00971.x (2007).

156 Ziprin, R. L., Young, C. R., Stanker, L. H., Hume, M. E. & Konkel, M. E. The absence of cecal colonization of chicks by a mutant of Campylobacter jejuni not expressing bacterial fibronectin-binding protein. *Avian Dis* **43**, 586-589 (1999).

157 Pei, Z. *et al.* Mutation in the peb1A locus of Campylobacter jejuni reduces interactions with epithelial cells and intestinal colonization of mice. *Infect Immun* **66**, 938-943 (1998).

158 Leon-Kempis Mdel, R., Guccione, E., Mulholland, F., Williamson, M. P. & Kelly, D. J. The Campylobacter jejuni PEB1a adhesin is an aspartate/glutamate-binding protein of an ABC transporter essential for microaerobic growth on dicarboxylic amino acids. *Mol Microbiol* **60**, 1262-1275, doi:10.1111/j.1365-2958.2006.05168.x (2006).

159 Young, K. T., Davis, L. M. & Dirita, V. J. Campylobacter jejuni: molecular biology and pathogenesis. *Nat Rev Microbiol* **5**, 665-679, doi:10.1038/nrmicro1718 (2007).

160 Pickett, C. L. *et al.* Prevalence of cytolethal distending toxin production in Campylobacter jejuni and relatedness of Campylobacter sp. cdtB gene. *Infect Immun* **64**, 2070-2078 (1996).

161 Lara-Tejero, M. & Galan, J. E. CdtA, CdtB, and CdtC form a tripartite complex that is required for cytolethal distending toxin activity. *Infect Immun* **69**, 4358-4365, doi:10.1128/IAI.69.7.4358-4365.2001 (2001).

162 Hu, X. & Stebbins, C. E. Dynamics and assembly of the cytolethal distending toxin. *Proteins* **65**, 843-855, doi:10.1002/prot.21167 (2006).

163 Lara-Tejero, M. & Galan, J. E. A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. *Science* **290**, 354-357 (2000).

164 Lara-Tejero, M. & Galan, J. E. Cytolethal distending toxin: limited damage as a strategy to modulate cellular functions. *Trends Microbiol* **10**, 147-152 (2002).

165 Purdy, D. *et al.* Characterisation of cytolethal distending toxin (CDT) mutants of Campylobacter jejuni. *J Med Microbiol* **49**, 473-479, doi:10.1099/0022-1317-49-5-473 (2000).

166 Mortensen, N. P. *et al.* The role of Campylobacter jejuni cytolethal distending toxin in gastroenteritis: toxin detection, antibody production, and clinical outcome. *APMIS* **119**, 626-634, doi:10.1111/j.1600-0463.2011.02781.x (2011).

167 Lertsethtakarn, P., Ottemann, K. M. & Hendrixson, D. R. Motility and chemotaxis in Campylobacter and Helicobacter. *Annu Rev Microbiol* **65**, 389-410, doi:10.1146/annurev-micro-090110-102908 (2011).

168 Yao, J. & Allen, C. Chemotaxis is required for virulence and competitive fitness of the bacterial wilt pathogen Ralstonia solanacearum. *J Bacteriol* **188**, 3697-3708, doi:10.1128/JB.188.10.3697-3708.2006 (2006).

169 O'Toole, R., Milton, D. L. & Wolf-Watz, H. Chemotactic motility is required for invasion of the host by the fish pathogen Vibrio anguillarum. *Mol Microbiol* **19**, 625-637 (1996).

170 Butler, S. M. & Camilli, A. Both chemotaxis and net motility greatly influence the infectivity of Vibrio cholerae. *Proc Natl Acad Sci U S A* **101**, 5018-5023, doi:10.1073/pnas.0308052101 (2004).

171 Wadhams, G. H. & Armitage, J. P. Making sense of it all: bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.* **5**, 1024-1037, doi:10.1038/nrm1524 (2004).

172 Sourjik, V. & Wingreen, N. S. Responding to chemical gradients: bacterial chemotaxis. *Curr Opin Cell Biol* **24**, 262-268, doi:10.1016/j.ceb.2011.11.008 (2012).

173 Bren, A. & Eisenbach, M. How signals are heard during bacterial chemotaxis: protein-protein interactions in sensory signal propagation. *J Bacteriol* **182**, 6865-6873 (2000).

174 Hazelbauer, G. L. & Lai, W. C. Bacterial chemoreceptors: providing enhanced features to two-component signaling. *Curr Opin Microbiol* **13**, 124-132, doi:10.1016/j.mib.2009.12.014 (2010).

175 Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends in biochemical sciences* **33**, 9-19, doi:10.1016/j.tibs.2007.09.014 (2008).

176 Wuichet, K. & Zhulin, I. B. Origins and diversification of a complex signal transduction system in prokaryotes. *Sci Signal* **3**, ra50, doi:10.1126/scisignal.2000724 (2010).

177 Zhao, R., Collins, E. J., Bourret, R. B. & Silversmith, R. E. Structure and catalytic mechanism of the E. coli chemotaxis phosphatase CheZ. *Nat Struct Biol* **9**, 570-575, doi:10.1038/nsb816 (2002).

178 Stock, J. B., Clarke, S. & Koshland, D. E., Jr. The protein carboxylmethyltransferase involved in Escherichia coli and Salmonella typhimurium chemotaxis. *Methods Enzymol* **106**, 310-321 (1984).

179 Marchant, J., Wren, B. & Ketley, J. Exploiting genome sequence: predictions for mechanisms of Campylobacter chemotaxis. *Trends Microbiol* **10**, 155-159 (2002).

180 Lowenthal, A. C. *et al.* A fixed-time diffusion analysis method determines that the three cheV genes of Helicobacter pylori differentially affect motility. *Microbiology* **155**, 1181-1191, doi:10.1099/mic.0.021857-0 (2009).

181 Hartley-Tassell, L. E. *et al.* Identification and characterization of the aspartate chemosensory receptor of *Campylobacter jejuni*. *Mol Microbiol* **75**, 710-730, doi:10.1111/j.1365-2958.2009.07010.x (2010).

182 Korolik, V. Aspartate chemosensory receptor signalling in Campylobacter jejuni. *Virulence* **1**, 414-417, doi:10.4161/viru.1.5.12735 (2010).

183 Alexander, R. P., Lowenthal, A. C., Harshey, R. M. & Ottemann, K. M. CheV: CheW-like coupling proteins at the core of the chemotaxis signaling network. *Trends Microbiol* **18**, 494-503, doi:10.1016/j.tim.2010.07.004 (2010).

184 Kanungpean, D., Kakuda, T. & Takai, S. Participation of CheR and CheB in the chemosensory response of Campylobacter jejuni. *Microbiology* **157**, 1279-1289, doi:10.1099/mic.0.047399-0 (2011).

185 Sweeney, E. G. & Guillemin, K. A gastric pathogen moves chemotaxis in a new direction. *MBio* **2**, doi:10.1128/mBio.00201-11 (2011).

186 Parkinson, J. S., Hazelbauer, G. L. & Falke, J. J. Signaling and sensory adaptation in Escherichia coli chemoreceptors: 2015 update. *Trends Microbiol* **23**, 257-266, doi:10.1016/j.tim.2015.03.003 (2015).

187 Alexander, R. P. & Zhulin, I. B. Evolutionary genomics reveals conserved structural determinants of signaling and adaptation in microbial chemoreceptors. *Proc Natl Acad Sci U S A* **104**, 2885-2890, doi:10.1073/pnas.0609359104 (2007).

188 Lacal, J., Garcia-Fontana, C., Munoz-Martinez, F., Ramos, J. L. & Krell, T. Sensing of environmental signals: classification of chemoreceptors according to the size of their ligand binding regions. *Environ. Microbiol.* **12**, 2873-2884, doi:10.1111/j.1462-2920.2010.02325.x (2010).

189 Briegel, A. *et al.* Structure of bacterial cytoplasmic chemoreceptor arrays and implications for chemotactic signaling. *Elife* **3**, e02151, doi:10.7554/eLife.02151 (2014).

190 Zhulin, I. B. The superfamily of chemotaxis transducers: from physiology to genomics and back. *Adv Microb Physiol* **45**, 157-198 (2001).

191 Wuichet, K., Alexander, R. P. & Zhulin, I. B. Comparative genomic and protein sequence analyses of a complex system controlling bacterial chemotaxis. *Methods in enzymology* **422**, 1-31, doi:10.1016/S0076-6879(06)22001-9 (2007).

192 Salah Ud-Din, A. I. & Roujeinikova, A. Methyl-accepting chemotaxis proteins: a core sensing element in prokaryotes and archaea. *Cell Mol Life Sci*, doi:10.1007/s00018-017-2514-0 (2017).

193 Ortega, A., Zhulin, I. B. & Krell, T. Sensory Repertoire of Bacterial Chemoreceptors. *Microbiol Mol Biol Rev* **81**, doi:10.1128/MMBR.00033-17 (2017).

194 Goy, M. F., Springer, M. S. & Adler, J. Sensory transduction in Escherichia coli: role of a protein methylation reaction in sensory adaptation. *Proc Natl Acad Sci U S A* **74**, 4964-4968 (1977).

195 Hazelbauer, G. L. Maltose chemoreceptor of Escherichia coli. *J Bacteriol* **122**, 206-214 (1975).

196 Gardina, P., Conway, C., Kossman, M. & Manson, M. Aspartate and maltose-binding protein interact with adjacent sites in the Tar chemotactic signal transducer of Escherichia coli. *J Bacteriol* **174**, 1528-1536 (1992).

197 Milburn, M. V. *et al.* Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* **254**, 1342-1347 (1991).

198 Mise, T. Structural Analysis of the Ligand-Binding Domain of the Aspartate Receptor Tar from Escherichia coli. *Biochemistry* **55**, 3708-3713, doi:10.1021/acs.biochem.6b00160 (2016).

199 Wu, H. *et al.* Identification and characterization of two chemotactic transducers for inorganic phosphate in Pseudomonas aeruginosa. *J Bacteriol* **182**, 3400-3404 (2000).

200 Luu, R. A. *et al.* Integration of chemotaxis, transport and catabolism in Pseudomonas putida and identification of the aromatic acid chemoreceptor PcaY. *Mol Microbiol* **96**, 134-147, doi:10.1111/mmi.12929 (2015).

201 Finn, R. D. *et al.* Pfam: clans, web tools and services. *Nucleic Acids Res* **34**, D247-251, doi:10.1093/nar/gkj149 (2006).

202 David, M. *et al.* Cascade regulation of nif gene expression in Rhizobium meliloti. *Cell* **54**, 671-683 (1988).

203 Christie, J. M. *et al.* Arabidopsis NPH1: a flavoprotein with the properties of a photoreceptor for phototropism. *Science* **282**, 1698-1701 (1998).

204 Moglich, A., Ayers, R. A. & Moffat, K. Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure* **17**, 1282-1294, doi:10.1016/j.str.2009.08.011 (2009).

205 Upadhyay, A. A., Fleetwood, A. D., Adebali, O., Finn, R. D. & Zhulin, I. B. Cache Domains That are Homologous to, but Different from PAS Domains Comprise the Largest Superfamily of Extracellular Sensors in Prokaryotes. *PLoS Comput. Biol.* **12**, e1004862; 1004810.1001371/journal.pcbi.1004862, doi:10.1371/journal.pcbi.1004862 (2016).

206 Bibikov, S. I., Biran, R., Rudd, K. E. & Parkinson, J. S. A signal transducer for aerotaxis in *Escherichia coli*. *Journal of bacteriology* **179**, 4075-4079 (1997).

207 Rebbapragada, A. *et al.* The Aer protein and the serine chemoreceptor Tsr independently sense intracellular energy levels and transduce oxygen, redox, and

energy signals for Escherichia coli behavior. *Proc Natl Acad Sci U S A* **94**, 10541-10546 (1997).

208 Sawai, H. *et al.* Structural basis for oxygen sensing and signal transduction of the heme-based sensor protein Aer2 from Pseudomonas aeruginosa. *Chem Commun (Camb)* **48**, 6523-6525, doi:10.1039/c2cc32549g (2012).

209 Garcia, D., Orillard, E., Johnson, M. S. & Watts, K. J. Gas Sensing and Signaling in the PAS-Heme Domain of the Pseudomonas aeruginosa Aer2 Receptor. *J Bacteriol* **199**, doi:10.1128/JB.00003-17 (2017).

210 Goers Sweeney, E. *et al.* Structure and proposed mechanism for the pH-sensing *Helicobacter pylori* chemoreceptor TlpB. *Structure* **20**, 1177-1188, doi:10.1016/j.str.2012.04.021 (2012).

211 Croxen, M. A., Sisson, G., Melano, R. & Hoffman, P. S. The *Helicobacter pylori* chemotaxis receptor TlpB (HP0103) is required for pH taxis and for colonization of the gastric mucosa. *Journal of bacteriology* **188**, 2656-2665, doi:10.1128/JB.188.7.2656-2665.2006 (2006).

212 Rader, B. A. *et al. Helicobacter pylori* perceives the quorum-sensing molecule AI-2 as a chemorepellent via the chemoreceptor TlpB. *Microbiology* **157**, 2445-2455, doi:10.1099/mic.0.049353-0 (2011).

213 Huang, J. Y. *et al.* Chemodetection and Destruction of Host Urea Allows *Helicobacter pylori* to Locate the Epithelium. *Cell Host Microbe* **18**, 147-156, doi:10.1016/j.chom.2015.07.002 (2015).

214 Anderson, J. K. *et al.* Chemorepulsion from the Quorum Signal Autoinducer-2 Promotes *Helicobacter pylori* Biofilm Dispersal. *MBio* **6**, e00379; 00310.01128/mBio.00379-00315, doi:10.1128/mBio.00379-15 (2015).

215 Garcia, V. *et al.* Identification of a Chemoreceptor for C2 and C3 Carboxylic Acids. *Appl. Environ. Microbiol.* **81**, 5449-5457, doi:10.1128/AEM.01529-15 (2015).

216 Nishiyama, S. *et al.* Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infection and immunity* **80**, 3170-3178, doi:10.1128/IAI.00039-12 (2012).

217 Nishiyama, S. *et al.* Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci. Rep.* **6**, 20866; 20810.21038/srep20866, doi:10.1038/srep20866 (2016).

218 Rico-Jimenez, M. *et al.* Paralogous chemoreceptors mediate chemotaxis towards protein amino acids and the non-protein amino acid gamma-aminobutyrate (GABA). *Mol. Microbiol.* **88**, 1230-1243, doi:10.1111/mmi.12255 (2013).

219 Glekas, G. D. *et al.* The *Bacillus subtilis* chemoreceptor McpC senses multiple ligands using two discrete mechanisms. *J. Biol. Chem.* **287**, 39412-39418, doi:10.1074/jbc.M112.413518 (2012).

220 Glekas, G. D. *et al.* A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. *J. Biol. Chem.* **285**, 1870-1878, doi:10.1074/jbc.M109.072108 (2010).

221 Liu, Y. C., Machuca, M. A., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. *Acta crystallographica. Section D, Biological crystallography* **71**, 2127-2136, doi:10.1107/S139900471501384X (2015).

222 McKellar, J. L., Minnell, J. J. & Gerth, M. L. A high-throughput screen for ligand binding reveals the specificities of three amino acid chemoreceptors from *Pseudomonas syringae* pv. *actinidiae*. *Mol. Microbiol.* **96**, 694-707, doi:10.1111/mmi.12964 (2015).

223 Webb, B. A., Hildreth, S., Helm, R. F. & Scharf, B. E. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Appl. Environ. Microbiol.* **80**, 3404-3415, doi:10.1128/AEM.00115-14 (2014).

224 Liu, X., Wood, P. L., Parales, J. V. & Parales, R. E. Chemotaxis to pyrimidines and identification of a cytosine chemoreceptor in *Pseudomonas putida*. *Journal of bacteriology* **191**, 2909-2916, doi:10.1128/JB.01708-08 (2009).

225 Fernandez, M., Morel, B., Corral-Lugo, A. & Krell, T. Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol. Microbiol.* **99**, 34-42, doi:10.1111/mmi.13215 (2016).

226 Machuca, M. A., Liu, Y. C., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni* Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition. *J. Struct. Biol.* **194**, 205-213, doi:10.1016/j.jsb.2016.02.019 (2016).

227 Pineda-Molina, E. *et al.* Evidence for chemoreceptors with bimodular ligand-binding regions harboring two signal-binding sites. *Proc. Natl. Acad. Sci. U S A* **109**, 18926-18931, doi:10.1073/pnas.1201400109 (2012).

228 Ortega, A. & Krell, T. The HBM domain: introducing bimodularity to bacterial sensing. *Protein Sci* **23**, 332-336, doi:10.1002/pro.2410 (2014).

229 Martin-Mora, D. *et al.* McpQ is a specific citrate chemoreceptor that responds preferentially to citrate/metal ion complexes. *Environ. Microbiol.* **18**, 3284-3295, doi:10.1111/1462-2920.13030 (2016).

230 Oku, S., Komatsu, A., Nakashimada, Y., Tajima, T. & Kato, J. Identification of Pseudomonas fluorescens chemotaxis sensory proteins for malate, succinate, and fumarate, and their involvement in root colonization. *Microbes Environ* **29**, 413-419, doi:10.1264/jsme2.ME14128 (2014).

231 Parales, R. E. *et al.* Pseudomonas putida F1 has multiple chemoreceptors with overlapping specificity for organic acids. *Microbiology* **159**, 1086-1096, doi:10.1099/mic.0.065698-0 (2013).

232 Draper, J., Karplus, K. & Ottemann, K. M. Identification of a chemoreceptor zinc-binding domain common to cytoplasmic bacterial chemoreceptors. *J Bacteriol* **193**, 4338-4345, doi:10.1128/JB.05140-11 (2011).

233 Collins, K. D., Lacal, J. & Ottemann, K. M. Internal sense of direction: sensing and signaling from cytoplasmic chemoreceptors. *Microbiol Mol Biol Rev* **78**, 672-684, doi:10.1128/MMBR.00033-14 (2014).

234 Schweinitzer, T. *et al.* Functional characterization and mutagenesis of the proposed behavioral sensor TlpD of *Helicobacter pylori*. *Journal of bacteriology* **190**, 3244-3255, doi:10.1128/JB.01940-07 (2008).

235 Collins, K. D. *et al.* The *Helicobacter pylori* CZB Cytoplasmic Chemoreceptor TlpD Forms an Autonomous Polar Chemotaxis Signaling Complex That Mediates a Tactic Response to Oxidative Stress. *Journal of bacteriology* **198**, 1563-1575, doi:10.1128/JB.00071-16 (2016).

236 Huang, J. Y., Goers Sweeney, E., Guillemin, K. & Amieva, M. R. Multiple Acid Sensors Control *Helicobacter pylori* Colonization of the Stomach. *PLoS Pathog.* **13**, e1006118; 1006110.1001371/journal.ppat.1006118, doi:10.1371/journal.ppat.1006118 (2017).

237 Li, Z. *et al.* Methyl-accepting chemotaxis proteins 3 and 4 are responsible for *Campylobacter jejuni* chemotaxis and jejuna colonization in mice in response to sodium deoxycholate. *J Med Microbiol* **63**, 343-354, doi:10.1099/jmm.0.068023-0 (2014).

238 Rahman, H. *et al.* Characterisation of a multi-ligand binding chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. *PLoS Pathog* **10**, e1003822, doi:10.1371/journal.ppat.1003822 (2014).

239 Mund, N. L. *et al.* Association of Campylobacter Jejuni ssp. Jejuni Chemotaxis Receptor Genes with Multilocus Sequence Types and Source of Isolation. *European journal of microbiology & immunology* **6**, 162-177, doi:10.1556/1886.2015.00041 (2016).

240 Zautner, A. E., Tareen, A. M., Gross, U. & Lugert, R. Chemotaxis in *Campylobacter jejuni*. *European journal of microbiology & immunology* **2**, 24-31, doi:10.1556/EuJMI.2.2012.1.5 (2012).

241 Day, C. J. *et al.* Variation of chemosensory receptor content of Campylobacter jejuni strains and modulation of receptor gene expression under different in vivo and in vitro growth conditions. *BMC Microbiol* **12**, 128, doi:10.1186/1471-2180-12-128 (2012).

242 Parkhill, J. *et al.* The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665-668, doi:10.1038/35001088 (2000).

243 Day, C. *et al.* Characterisation of the galactose chemosensory receptor of *Campylobacter jejuni*. *In: Proceedings of The Australian Society for Microbiology Annual Scientific Meeting. Australian Society for Microbiology.* **Abstract #107.** (2014).

244 Reuter, M. & van Vliet, A. H. Signal balancing by the CetABC and CetZ chemoreceptors controls energy taxis in Campylobacter jejuni. *PLoS One* **8**, e54390, doi:10.1371/journal.pone.0054390 (2013).

245 Tareen, A. M., Dasti, J. I., Zautner, A. E., Gross, U. & Lugert, R. *Campylobacter jejuni* proteins Cj0952c and Cj0951c affect chemotactic behaviour towards formic acid and are important for invasion of host cells. *Microbiology* **156**, 3123-3135, doi:10.1099/mic.0.039438-0 (2010).

246 Zautner, A. E. *et al.* Epidemiological association of different Campylobacter jejuni groups with metabolism-associated genetic markers. *Appl Environ Microbiol* **77**, 2359-2365, doi:10.1128/AEM.02403-10 (2011).

247 Dasti, J. I., Tareen, A. M., Lugert, R., Zautner, A. E. & Gross, U. *Campylobacter jejuni*: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. *Int J Med Microbiol* **300**, 205-211, doi:10.1016/j.ijmm.2009.07.002 (2010).

248 Kim, K. K., Yokota, H. & Kim, S. H. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* **400**, 787-792, doi:10.1038/23512 (1999).

249 Cerda, O. A. *et al.* tlpA gene expression is required for arginine and bicarbonate chemotaxis in Helicobacter pylori. *Biological research* **44**, 277-282, doi:/S0716-97602011000300009 (2011).

Chapter 2: Structural basis for ligand recognition by the dCache ligand binding domain of *C. jejuni* transducer-like protein 3 (Tlp3)

Preface for Chapter 2

The molecular mechanisms involved in *C. jejuni* pathogenesis remain poorly understood. However, several virulence factors including flagellar motility and chemotaxis are known to play a major role during the colonisation and infection of both avian and mammalian hosts [1-3]. *C. jejuni* responds chemotactically to various environmental stimuli [4] sensed by a set of membrane-associated chemoreceptors (group A Tlps) [5,6]. Tlps members of group A are predicted to comprise a periplasmic LBD flanked by two transmembrane helices and a highly conserved cytoplasmic signalling domain.

To date, the signal specificity of some *C. jejuni* chemoreceptors has been identified. Tlp3 [also termed as CcmL (*Campylobacter* chemoreceptor for multiple ligands) or Cj1564] receptor has been shown to recognise several amino acids, tricarboxylic acid cycle intermediates, bile acids and nucleotides [7,8]. Interestingly, Tlp3 was found to mediate both chemoattractant and chemorepellent responses. For example, *tlp3* isogenic mutants were shown to have a reduction of the chemotactic response to chemoattractants, such as: isoleucine, purine, malic acid, fumaric acid, aspartic acid and sodium deoxycholate [7,8]; as well as the repellents lysine, glucosamine, succinic acid, arginine and thiamine [7].

Tlp3-LBD shares no significant sequence similarity to chemoreceptors from non-*Epsilonproteobacteria* [7,8]. Despite its low sequence identity, a search for domain architecture in the Pfam database [9] detected the presence of the conserved dCache_1 motif [10] in Tlp3 sequence. The dCache_1 motif has been found in chemoreceptors [11,12] and members from the family 1 histidine kinases [13], of which the sensing module has two tandem PAS-like subdomains. Thus, it was hypothesised that the Tlp3-LBD has similar structural features. The following experiments aimed to determine the structural basis of Tlp3 ligand recognition and to identify the structural changes this receptor undergoes upon ligand binding.

This chapter contains a brief report published in January 2015 in the *Acta Crystallographica Section F* describing the cloning, overexpression, refolding, purification, crystallisation and preliminary analysis of the X-ray diffraction data [14]. The detailed biophysical, biochemical and structural analysis is presented in the journal article format and was published in July 2015 in the *Acta Crystallographica Section D* [15].

Monash University

Declaration for Thesis Chapter 2 (Section 2.1)

Declaration by candidate

In the case of Chapter 2 (section 2.1), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiments, performed experiments, analysed data and wrote manuscript	48

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Yu C. Liu	Designed experiments, analysed data and reviewed manuscript	
Simone A. Beckham	Analysed data, provided intellectual input and reviewed manuscript	
Anna Roujeinikova	Designed experiments, analysed data, wrote manuscript and led the research	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's Signature	Mayra A. Machuca Pérez	**Date** 10/12/17
Main Supervisor's Signature		**Date** 13 /12/17

Section 2.1.

CrossMark

research communications

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Cloning, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for multiple ligands (CcmL)

Mayra A. Machuca,[a] Yu C. Liu,[a] Simone A. Beckham[b] and Anna Roujeinikova[a]*

[a]Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia, and [b]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia. *Correspondence e-mail: anna.roujeinikova@monash.edu

Received 15 November 2014
Accepted 9 January 2015

Keywords: *Campylobacter jejuni*; chemotaxis; transducer-like proteins; methyl-accepting proteins.

A periplasmic sensory domain of the *Campylobacter jejuni* chemoreceptor for multiple ligands (CcmL) has been crystallized by the hanging-drop vapour-diffusion method using polyethylene glycol 3350 as a precipitating agent. A complete data set was collected to 1.3 Å resolution using cryocooling conditions and synchrotron radiation. The crystals belonged to space group $P2_1$, with unit-cell parameters a = 42.6, b = 138.0, c = 49.0 Å, β = 94.3°.

1. Introduction

Campylobacter jejuni is one of the most common causes of human bacterial gastroenteritis (Allos, 2001; Butzler, 2004). It colonizes the gastrointestinal tract of poultry and cattle and is transmitted to humans *via* contaminated food (predominantly chicken products) or water (Allos, 2001; Butzler, 2004; Inglis *et al.*, 2005; Wingstrand *et al.*, 2006). *C. jejuni* infections cause fever, abdominal pain and mild to severe diarrhoea (Allos, 2001; Butzler, 2004; Wassenaar & Blaser, 1999). Post-infection complications may include arthritis, reactive myositis and Guillain–Barré syndrome (Schmidt-Ott *et al.*, 2006; Allos, 2001; Friedman *et al.*, 2000). Although the molecular mechanisms involved in its pathogenesis remain poorly understood, several virulence factors, such as chemotaxis, adherence and lipooligosaccharide, have been identified (Young *et al.*, 2007).

Chemotaxis allows motile bacteria to swim in a certain direction in response to extracellular chemical signals. It has been shown to play an important role in initial colonization and development of the disease in animal models (Josenhans & Suerbaum, 2002). *C. jejuni* is attracted to amino acids, organic acids and carbohydrates (Rahman *et al.*, 2014; Hartley-Tassell *et al.*, 2010; Hugdahl *et al.*, 1988). The bacteria sense these chemicals *via* transducer-like proteins (Tlps) located in the bacterial membrane or cytoplasm (Zautner *et al.*, 2012). Interaction between the Tlps and their respective ligands initiates a molecular signal transduction cascade that causes a change in the direction of rotation of flagellar motors.

11 putative Tlps have been identified in *C. jejuni* (Day *et al.*, 2012; Zautner *et al.*, 2012; Parkhill *et al.*, 2000); these chemoreceptors have been classified into three groups according to their predicted domain organization. Tlp1, Tlp2, Tlp3, Tlp4, Tlp7$_{mc}$, Tlp7$_{m}$, Tlp10 and Tlp11 belong to group A, Tlp9 belongs to group B and Tlp5, Tlp6, Tlp7$_{c}$ and Tlp8 belong to group C (Zautner *et al.*, 2012). Group A Tlps have a similar structure to the methyl-accepting proteins (MCPs) found in

Escherichia coli and to family A transducers of *Halobacterium salinarum* (Zhang *et al.*, 1996). They are predicted to contain an N-terminal transmembrane helix followed by a periplasmic ligand-binding domain, a second transmembrane helix and a C-terminal cytoplasmic signalling domain (Marchant *et al.*, 2002; Zautner *et al.*, 2012). While the signalling domain is highly conserved (and is even identical in the case of Tlp2, Tlp3 and Tlp4), the periplasmic domain exhibits high variability between different Tlps (Parkhill *et al.*, 2000; Marchant *et al.*, 2002).

Four of the group A Tlp receptors of *C. jejuni* (Tlp1, Tlp3, Tlp7 and Tlp11) have been characterized to date. The aspartate receptor Tlp1, also termed CcaA (*Campylobacter* chemoreceptor for aspartate A; Hartley-Tassell *et al.*, 2010), the multiligand-binding receptor Tlp3, also termed CcmL (*Campylobacter* chemoreceptor for multiple ligands; Rahman *et al.*, 2014), and the formic acid receptor Tlp7 (Tareen *et al.*, 2010) are present in most sequenced strains of *C. jejuni*, whereas the galactose receptor Tlp11 is the least common and has only been found in clinical isolates from patients hospitalized owing to *C. jejuni* infection (Day *et al.*, 2012, 2014).

The sensory domain of CcmL, which has been shown to bind multiple ligands including isoleucine, lysine, glucosamine, aspartate, succinic acid, arginine, malic acid, thiamine and sodium deoxycholate, shares no significant sequence similarity to chemoreceptors from non-Epsilonproteobacteria (Rahman *et al.*, 2014; Li *et al.*, 2014). This suggests that the mechanism of ligand recognition by CcmL may be unique to *Campylobacter* spp. and *Campylobacter*-related organisms. Furthermore, a *BLAST* sequence-similarity search against structures deposited in the Protein Data Bank (PDB) identified no close structural homologues of this domain. However, a Pfam analysis (Finn *et al.*, 2014) of the CcmL primary sequence revealed the presence of a Cache_1 (**ca**lcium channels and **che**motaxis receptors) motif located between residues 163 and 237 (Anantharaman & Aravind, 2000; Rahman *et al.*, 2014). The Cache motif is a signature of double-PAS sensor domains found in family 1 histidine kinases (Zhang & Hendrickson, 2010) and chemotaxis receptors (Nishiyama *et al.*, 2012). In agreement with this, a search against the PDB using pairwise comparison of profile hidden Markov models implemented on the *HHpred* server (http://toolkit.tuebingen.mpg.de/hhpred#; Söding *et al.*, 2005) identified remote homology at the level of secondary structure (but not sequence) between the sensory domain of CcmL and the double-PAS periplasmic domains of the putative *Vibrio cholerae* chemoreceptor Mlp37 (PDB entry 3c8c; New York SGX Research Center for Structural Genomics, unpublished work) and family 1 histidine kinases (PDB entries 3lib and 3li9; Zhang & Hendrickson, 2010). This permits tentative assignment of the periplasmic domain of CcmL to the recently described novel family of tandem-PAS chemoreceptor sensing domains (Glekas *et al.*, 2010; Nishiyama *et al.*, 2012).

We have recently reported the crystallization of the periplasmic sensory domain of Tlp1 (CcaA) as part of a study to identify the structural features that determine its ligand specificity (Machuca *et al.*, 2015). Tlp1 does not appear to have a Cache motif. In this paper, we report the expression, refolding from inclusion bodies, purification and crystallization of the periplasmic domain of CcmL (amino-acid residues 42–291). The determination of the first crystal structure of a double-PAS sensory domain from a characterized chemoreceptor will help us to understand the structural basis of ligand recognition in this novel class of receptor-sensing domains.

Table 1
Macromolecule-production information.

Source organism	*C. jejuni* serotype O:2 (strain NCTC 11168)
DNA source	Plasmid pGU0816 (Rahman *et al.*, 2014)
Forward primer	CACCAAAACCTCACTATATGAAAGCAC
Reverse primer	CTCGAGTTAAGCTTTATAAATAGGTTTATTTATA-AT
Cloning vector	pET151/D-TOPO
Expression vector	pET151/D-TOPO
Expression host	*E. coli* strain BL21-CodonPlus (DE3)-RIPL
Complete amino-acid sequence of the construct produced†	MHHHHHHGKPIPNPLLGLDSTENLYFQgidpftK-TSLYESTLKNQTDLLKVTQSTVEDFRSTNQSF-TRALEKDIANLPYQSLITEENIINNVGPILKY-YRHSINALNVYLGLNNGKVLLSQKSNDAKMPE-LRDDLDIKTKDWYQEALKTNDIFVTPAYLDTV-LKQYVITYSKAIYKDGKIIGVLGVDIPSEDLQ-NLVAKTPGNTFLFDQKNKIFAATNKELLNPSI-DHSPVLNAYKLNGDNNFFSYKLNNEERLGACT-KVFAYTACITESADIINKPIYKA

† The TEV cleavage site is underlined and the additional sequence at the N-terminus (cloning artifact) is shown in lower case.

2. Materials and methods

2.1. Cloning, protein expression, solubilization of inclusion bodies, protein refolding and purification

The DNA fragment encoding the periplasmic sensory domain of CcmL ($CcmL^{peri}$; amino-acid residues 42–291; UniProtKB Q0P864; also known as Tlp3 or Cj1564) was amplified from plasmid pGU0816 (Rahman *et al.*, 2014). The PCR reaction was carried out using *Pfu* DNA polymerase (Stratagene) and the primers CACCAAAACCTCACTATATGAAAGCAC (forward) and CTCGAGTTAAGCTTTATAAATAGGTTTATTTATAAT (reverse). The amplified fragment was ligated into the pET151/D-TOPO vector using the TOPO cloning kit (Invitrogen) to produce an expression construct that contained an N-terminal His_6 tag. The correct insertion of the fragment was confirmed by DNA sequencing.

For recombinant protein expression, *E. coli* BL21-CodonPlus(DE3)-RIPL (Stratagene) cells were transformed with the vector and grown at 37°C in LB medium containing 50 µg ml^{-1} ampicillin until the $OD_{600\ nm}$ reached 0.6, at which point expression was induced by adding 1 m*M* isopropyl β-D-1-thiogalactopyranoside. The cells were grown for a further 4 h and harvested by centrifugation at 6000*g* for 15 min at 4°C. The cell pellet was resuspended in buffer *A* consisting of 10 m*M* Tris–HCl pH 8.0 and lysed using an Avestin EmulsiFlex-C5 high-pressure homogenizer. The lysate was centrifuged at 10 000*g* for 15 min at 4°C. Analysis of the supernatant and pellet using SDS–PAGE showed that the recombinant

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Temperature (K)	293
Protein concentration (mg ml^{-1})	15
Composition of reservoir solution	19%(*w*/*v*) polyethylene glycol 3350, 100 m*M* sodium citrate pH 5.0, 200 m*M* ammonium sulfate

protein was predominantly deposited in inclusion bodies. Details of macromolecule production are summarized in Table 1.

To isolate CcmL^{peri} from inclusion bodies, we used a refolding procedure that we recently developed for the production of a soluble, crystallizable form of the periplasmic sensory domain of the bacterial chemoreceptor TlpC (Liu & Roujeinikova, 2015), with minor modifications. The inclusion-bodies pellet was washed three times with buffer *B* [10 m*M* Tris–HCl pH 8.0, 0.2 m*M* phenylmethanesulfonyl fluoride (PMSF), 1% Triton X-100] and once in buffer *C* (10 m*M* Tris–HCl pH 8.0, 0.2 m*M* PMSF). The inclusion bodies were solubilized in buffer *D* [10 m*M* Tris–HCl pH 8.0, 8 *M* urea, 10 m*M* dithiothreitol (DTT), 0.2 m*M* PMSF] for 30 min at 4°C with mixing by axial rotation. The mixture was then clarified by centrifugation at 30 000*g* for 30 min at 4°C and the protein concentration in the supernatant was measured using the Bradford assay (Bradford, 1976). 60 mg of denatured protein was refolded by dilution into 250 ml buffer *E* (3 *M* urea, 100 m*M* Tris–HCl pH 8.0, 0.4 *M* L-arginine monohydrochloride, 20 m*M* reduced L-glutathione, 2 m*M* oxidized L-glutathione) followed by a 48 h incubation at 4°C with continuous mixing. The sample was then dialyzed against 5 l buffer *A* for 16 h at 4°C. NaCl, Tris–HCl pH 8.0 and imidazole were added to the protein solution to final concentrations of 500, 20 and 20 m*M*, respectively, and the sample was loaded onto a 5 ml HiTrap Chelating HP column (GE Healthcare) pre-washed with buffer *F* (500 m*M* NaCl, 20 m*M* Tris–HCl pH 8.0, 20 m*M* imidazole). The column was washed with 30 column volumes of the same buffer and the protein was eluted with buffer *G* (500 m*M* NaCl, 20 m*M* Tris–HCl pH 8.0 and 500 m*M* imidazole).

The tag was cleaved off with His_6-TEV protease (Invitrogen) whilst dialysing the sample against buffer *H* [150 m*M* NaCl, 10 m*M* Tris–HCl pH 8.0, 2 m*M* DTT, 1%(*v*/*v*) glycerol]. NaCl and imidazole were then added to the sample to final concentrations of 500 and 20 m*M*, respectively, and the TEV protease and uncleaved protein were removed on a HiTrap Chelating HP column. The flowthrough was concentrated to 400 µl in an Amicon Ultra Ultracel 10 kDa cutoff concentrator and passed through a Superdex 200 HiLoad 26/60 gel-filtration column (GE Healthcare) equilibrated with buffer *I* (10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl). The protein purity was judged to be greater than 90% (Fig. 1).

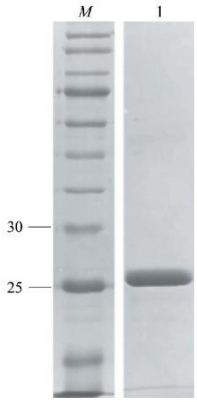


Figure 1
Coomassie Blue-stained 12% SDS–PAGE gel of pure, untagged recombinant CcmL^{peri} used for the crystallization and SEC-MALS experiments. Lane *M* contains molecular-weight marker (labelled in kDa).

2.2. Crystallization

CcmL^{peri} was concentrated to 15 mg ml^{-1} in buffer *I*, incubated for 30 min with 8 m*M* isoleucine and centrifuged for 20 min at 13 000*g* to clarify the solution. Initial screening of crystallization conditions was carried out by the hanging-drop vapour-diffusion method using a Phoenix crystallization robot (Art Robbins Instruments) and the screens JBScreen HTS I and II (Jena Bioscience), The JCSG+ Suite (Qiagen), Crystal Screen HT and PEG/Ion HT (Hampton Research). The droplets in the initial crystallization trials consisted of 200 nl protein solution mixed with 200 nl reservoir solution and were equilibrated against 50 µl reservoir solution in a 96-well Art Robbins CrystalMation Intelli-Plate (Hampton Research). Crystals appeared after 1 d from condition E4 of JBScreen HTS I (Fig. 2), which consisted of 25%(*w*/*v*) polyethylene glycol 4000, 100 m*M* sodium citrate pH 5.6, 200 m*M* ammonium sulfate. Subsequent refinement of this condition using 2 µl hanging drops suspended over 500 µl reservoir solution

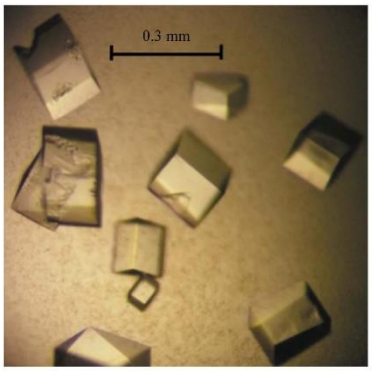


Figure 2
Crystals of a putative CcmL^{peri} complex with isoleucine.

yielded an optimal crystallization reservoir solution consisting of 19%(*w*/*v*) polyethylene glycol 3350, 100 m*M* sodium citrate pH 5.0, 200 m*M* ammonium sulfate (Table 2).

2.3. X-ray diffraction data collection and processing and size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis

The CcmL^{peri} crystal was washed for 1 s in a cryoprotectant solution consisting of 24%(*w*/*v*) polyethylene glycol 3350, 100 m*M* sodium citrate pH 5.0, 200 m*M* ammonium sulfate, 10%(*v*/*v*) glycerol, 10 m*M* isoleucine and was flash-cooled by plunging it into liquid nitrogen. X-ray data were collected to 1.3 Å resolution on the MX2 beamline at the Australian Synchrotron (Fig. 3). A total of 360 images were collected using a 0.5° oscillation width. The data were processed using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (Evans & Murshudov, 2013) from the *CCP*4 suite (Winn *et al.*, 2011). Calculation of the self-rotation function was performed using *POLARRFN* (Winn *et al.*, 2011). Statistics for data collection and processing are presented in Table 3.

The absolute hydrated molecular mass of CcmL^{peri} in solution was determined by MALS analysis coupled to size-exclusion chromatography (SEC). CcmL^{peri} was dialysed against buffer *J* (100 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl) and concentrated to 100 µ*M*. A 100 µl sample was loaded onto a WTC-030S5 size-exclusion column (Wyatt) pre-equilibrated with buffer *J* flowing at 0.4 ml min^{-1}. The eluate was passed through an in-line DAWN HELEOS light-scattering detector, an Optilab T-rEX differential refractive-index detector (Wyatt) and a quasi-elastic light-scattering detector (Wyatt-QELS, Wyatt). A bovine serum albumin (BSA) standard was used to normalize the MALS detectors. For calculations of the molecular mass and hydrodynamic radius, the light-scattered intensity and the refractive index were analyzed using *ASTRA* 6.0 (Wyatt), with a value for the refractive-index increment (d*n*/dc) for the protein of 0.185 ml g^{-1}.

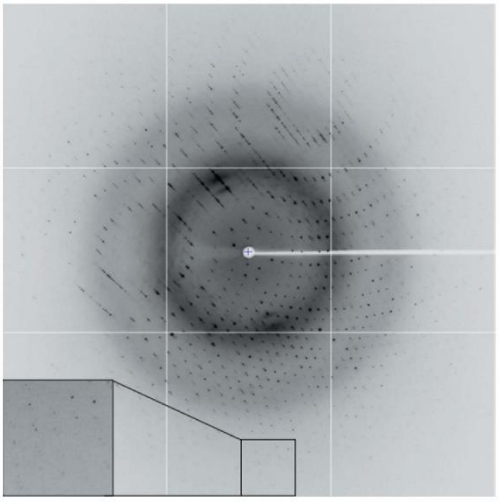


Figure 3
A representative 0.5° oscillation image of the data collected using an ADSC Quantum 315r CCD detector at station MX2 at the Australian Synchrotron, Victoria, Australia. The magnified rectangle shows diffraction spots between 1.4 and 1.6 Å resolution.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	MX2 beamline, Australian Synchrotron
Wavelength (Å)	0.95
Temperature (K)	100
Detector	ADSC Quantum 315r CCD
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	1
Space group	$P2_1$
a, *b*, *c* (Å)	42.6, 138.0, 49.0
α, β, γ (°)	90.0, 94.3, 90.0
Mosaicity (°)	0.4
Resolution range (Å)	40.0–1.3 (1.32–1.30)
Total No. of reflections	441329 (10462)
No. of unique reflections	128260 (4046)
Completeness (%)	93 (59)
Multiplicity	3.4 (2.6)
$\langle I/\sigma(I)\rangle$	12.2 (2.1)
$R_{r.i.m.}$	0.045 (0.295)
Overall *B* factor from Wilson plot (Å^2)	11.9

3. Results and discussion

Recombinant CcmL^{peri} was expressed with a cleavable N-terminal His$_6$ tag from the pET151/D-TOPO plasmid in *E. coli* BL21(DE3)-RIPL cells upon induction of T7 polymerase. The expression predominantly resulted in protein deposition in inclusion bodies (approximately 100 mg CcmL^{peri}-His$_6$ in the form of washed inclusion bodies per litre of culture). CcmL^{peri}-His$_6$ was isolated from inclusion bodies by following the refolding procedure for the production of a soluble, crystallizable form of periplasmic sensory domains of bacterial chemoreceptors that we have recently reported (Liu & Roujeinikova, 2015). Refolding and purification of protein from the inclusion bodies (including tag removal) resulted in >90% electrophoretic homogeneity based on Coomassie Blue staining of SDS–PAGE gels (Fig. 1). Approximately 3 mg pure untagged CcmL^{peri} was obtained per litre of culture. The recombinant protein used for the preliminary analysis and crystallization comprised amino-acid residues 42–291 of CcmL plus six additional residues from the TEV cleavage site and the vector (GIDPFT). The protein migrated on SDS–PAGE with an apparent molecular weight of 26 kDa, which is close to the value calculated from the amino-acid sequence (28.7 kDa).

To determine the oligomeric state of CcmL^{peri} in solution, multi-angle light-scattering (MALS) analysis coupled to size-exclusion chromatography was carried out. Both in the presence and in the absence of 10 m*M* isoleucine, CcmL^{peri} eluted as a single symmetrical peak with a polydispersity index value close to 1 (data not shown), indicating that the eluted

particles were homogenous with respect to molar mass. The elution volumes with and without isoleucine were almost identical. The derived molecular weight was 27 kDa, indicating that the periplasmic sensory domain of CcmL is monomeric under the tested buffer conditions. Although methyl-accepting chemoreceptor proteins function as dimers and often possess a dimeric sensory domain (Hazelbauer *et al.*, 2008), we note that, like $CcaA^{peri}$, $CcmL^{peri}$ shows no significant sequence similarity to any of the characterized chemoreceptors and hypothesize that dimerization of CcaA and CcmL may be primarily driven by dimerization of the transmembrane and cytoplasmic domains. There is a different reported example of a sensory domain of a dimeric methyl-accepting chemoreceptor (specifically, a *Salmonella typhimurium* aspartate chemoreceptor) that has been observed in solution as both monomers and dimers, with dimerization being dependent on protein concentration (Milligan & Koshland, 1993).

Crystals of the putative $CcmL^{peri}$ complex with isoleucine (Fig. 2) were obtained using a sparse-matrix crystallization approach. Analysis of the test data collected from a cryo-cooled $CcmL^{peri}$ crystal using the Australian Synchrotron indicated that these crystals diffracted to at least 1.3 Å resolution (Table 3). Autoindexing using *iMosflm* showed that they belonged to a primitive monoclinic space group, with unit-cell parameters a = 42.6, b = 138.0, c = 49.0 Å, β = 94.3°. Data analysis using *POINTLESS* from the *CCP*4 suite showed systematic absences along the $0k0$ axis, with only reflections with $k = 2n$ present, which suggested that the crystals belonged to space group $P2_1$. Calculations of the Matthews coefficient for two molecules in the asymmetric unit gave a value of 2.5 $Å^3$ Da^{-1}, which lies in the range observed for protein crystals (Matthews, 1977; Kantardjieff & Rupp, 2003). In a self-rotation function calculated using data in the resolution range 10–6 Å with an integration radius of 25 Å, no dominant features that can be confidently assigned to noncrystallographic axes were found in the κ = 90, 120 or 180° sections. Thus, we are currently unable to determine whether the asymmetric unit contains a symmetric dimer or two independent monomers. A search for heavy-atom derivatives and the determination of the structure using the single- or multiple-wavelength anomalous dispersion method is under way.

Acknowledgements

Plasmid pGU0816 was kindly provided by Professor Victoria Korolik, Griffith University, Australia. We thank the staff at the Australian Synchrotron for their assistance with data collection. We also thank Dr Danuta Maksel and Dr Robyn Gray at the Monash Crystallography Unit for assistance in setting up robotic crystallization trials. This work was funded by the Australian Research Council Research Fellowship to AR.

References

Allos, B. M. (2001). *Clin. Infect. Dis.* **32**, 1201–1206.
Anantharaman, V. & Aravind, L. (2000). *Trends Biochem. Sci.* **25**, 535–537.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Butzler, J. P. (2004). *Clin. Microbiol. Infect.* **10**, 868–876.
Day, C. J., Hartley-Tassell, L. E., Shewell, L. K., King, R. M., Tram, G., Day, S. K., Semchenko, E. A. & Korolik, V. (2012). *BMC Microbiol.* **12**, 128.
Day, C., King, R., Tram, G., Hartley-Tassell, L., Fleetwood, A., Zhulin, I. & Korolik, V. (2014). *Proceedings of The Australian Society for Microbiology Annual Scientific Meeting 2014*, Abstract #107. Fitzroy: Australian Society for Microbiology. http://asmicro-2014.p.asnevents.com.au/days/2014-07-08/abstract/13760.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J. & Punta, M. (2014). *Nucleic Acids Res.* **42**, D222–D230.
Friedman, C., Neimann, J., Wegener, H. & Tauxe, R. (2000). *Campylobacter*, edited by I. Nachamkin & M. J. Blaser, pp. 121–138. Washington: ASM Press.
Glekas, G. D., Foster, R. M., Cates, J. R., Estrella, J. A., Wawrzyniak, M. J., Rao, C. V. & Ordal, G. W. (2010). *J. Biol. Chem.* **285**, 1870–1878.
Hartley-Tassell, L. E., Shewell, L. K., Day, C. J., Wilson, J. C., Sandhu, R., Ketley, J. M. & Korolik, V. (2010). *Mol. Microbiol.* **75**, 710–730.
Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. (2008). *Trends Biochem. Sci.* **33**, 9–19.
Hugdahl, M. B., Beery, J. T. & Doyle, M. P. (1988). *Infect. Immun.* **56**, 1560–1566.
Inglis, G. D., Kalischuk, L. D., Busz, H. W. & Kastelic, J. P. (2005). *Appl. Environ. Microbiol.* **71**, 5145–5153.
Josenhans, C. & Suerbaum, S. (2002). *Int. J. Med. Microbiol.* **291**, 605–614.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Li, Z., Lou, H., Ojcius, D. M., Sun, A., Sun, D., Zhao, J., Lin, X. & Yan, J. (2014). *J. Med. Microbiol.* **63**, 343–354.
Liu, Y. C. & Roujeinikova, A. R. (2015). *Protein Expr. Purif.* **107**, 29–34.
Machuca, M. A., Liu, Y. C. & Roujeinikova, A. (2015). *Acta Cryst.* F**71**, 110–113.
Marchant, J., Wren, B. & Ketley, J. (2002). *Trends Microbiol.* **10**, 155–159.
Matthews, B. W. (1977). *The Proteins*, 3rd ed., edited by H. Neurath & R. L. Hill, Vol. 3, pp. 468–477. New York: Academic Press.
Milligan, D. L. & Koshland, D. E. Jr (1993). *J. Biol. Chem.* **268**, 19991–19997.
Nishiyama, S.-I., Suzuki, D., Itoh, Y., Suzuki, K., Tajima, H., Hyakutake, A., Homma, M., Butler-Wu, S. M., Camilli, A. & Kawagishi, I. (2012). *Infect. Immun.* **80**, 3170–3178.
Parkhill, J. *et al.* (2000). *Nature (London)*, **403**, 665–668.
Rahman, H., King, R. M., Shewell, L. K., Semchenko, E. A., Hartley-Tassell, L. E., Wilson, J. C., Day, C. J. & Korolik, V. (2014). *PLoS Pathog.* **10**, e1003822.
Schmidt-Ott, R., Schmidt, H., Feldmann, S., Brass, F., Krone, B. & Gross, U. (2006). *Clin. Vaccine Immunol.* **13**, 779–783.
Söding, J., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
Tareen, A. M., Dasti, J. I., Zautner, A. E., Gross, U. & Lugert, R. (2010). *Microbiology*, **156**, 3123–3135.
Wassenaar, T. M. & Blaser, M. J. (1999). *Microbes Infect.* **1**, 1023–1033.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Wingstrand, A., Neimann, J., Engberg, J., Nielsen, E. M., Gerner-Smidt, P., Wegener, H. C. & Mølbak, K. (2006). *Emerg. Infect. Dis.* **12**, 280–284.
Young, K. T., Davis, L. M. & DiRita, V. J. (2007). *Nature Rev. Microbiol.* **5**, 665–679.

Zautner, A. E., Tareen, A. M., Gross, U. & Lugert, R. (2012). *Eur. J. Microbiol. Immunol.* **2**, 24–31.
Zhang, W., Brooun, A., McCandless, J., Banda, P. & Alam, M. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 4649–4654.
Zhang, Z. & Hendrickson, W. A. (2010). *J. Mol. Biol.* **400**, 335–353.

Monash University

Declaration for Thesis Chapter 2 (Section 2.2)

Declaration by candidate

In the case of Chapter 2 (Section 2.2), the nature and extent of my contribution to the work was the following:

Nature of contribution	**Extent of contribution (%)**
Designed experiments, performed experiments, analysed data and wrote manuscript	41

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	**Nature of contribution**	**Extent of contribution (%) for student co-authors only**
Yu C. Liu	Designed experiments, analysed data and reviewed manuscript	
Simone A. Beckham	Analysed data, provided intellectual input and reviewed manuscript	
Menachem J. Gunzburgb	Analysed data, provided intellectual input and reviewed manuscript	
Anna Roujeinikova	Designed experiments, analysed data, wrote manuscript and led the research	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's Signature	Mayra A. Machuca Pérez	**Date** 10/12/17
Main Supervisor's Signature		**Date** 13 /12/17

Section 2.2.

CrossMark

electronic reprint

BIOLOGICAL
CRYSTALLOGRAPHY

ISSN: 1399-0047
journals.iucr.org/d

Structural basis for amino-acid recognition and transmembrane signalling by tandem Per–Arnt–Sim (tandem PAS) chemoreceptor sensory domains

Yu C. Liu, Mayra A. Machuca, Simone A. Beckham, Menachem J. Gunzburg and Anna Roujeinikova

Acta Cryst. (2015). D**71**, 2127–2136

BIOLOGICAL CRYSTALLOGRAPHY
ISSN 1399-0047

Structural basis for amino-acid recognition and transmembrane signalling by tandem Per–Arnt–Sim (tandem PAS) chemoreceptor sensory domains

Yu C. Liu,[a]‡ Mayra A. Machuca,[a]‡ Simone A. Beckham,[b] Menachem J. Gunzburg[b] and Anna Roujeinikova[a,b]*

[a]Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia, and [b]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia. *Correspondence e-mail: anna.roujeinikova@monash.edu

Received 3 June 2015
Accepted 21 July 2015

Edited by Z. Dauter, Argonne National Laboratory, USA

‡ YCL and MPM contributed equally to this work.

Keywords: chemotaxis; methyl-accepting protein; sensing domain.

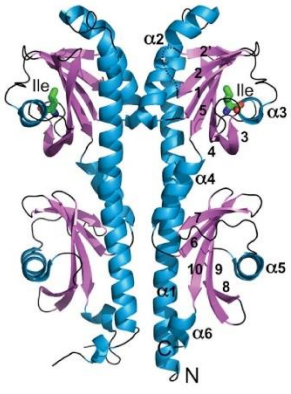
PDB references: Tlp3 PTPSD, 4xmq; complex with isoleucine, 4xmr

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

Chemotaxis, mediated by methyl-accepting chemotaxis protein (MCP) receptors, plays an important role in the ecology of bacterial populations. This paper presents the first crystallographic analysis of the structure and ligand-induced conformational changes of the periplasmic tandem Per-Arnt-Sim (PAS) sensing domain (PTPSD) of a characterized MCP chemoreceptor. Analysis of the complex of the *Campylobacter jejuni* Tlp3 PTPSD with isoleucine (a chemoattractant) revealed that the PTPSD is a dimer in the crystal. The two ligand-binding sites are located in the membrane-distal PAS domains on the faces opposite to the dimer interface. Mutagenesis experiments show that the five strongly conserved residues that stabilize the main-chain moiety of isoleucine are essential for binding, suggesting that the mechanism by which this family of chemoreceptors recognizes amino acids is highly conserved. Although the fold and mode of ligand binding of the PTPSD are different from the aspartic acid receptor Tar, the structural analysis suggests that the PTPSDs of amino-acid chemoreceptors are also likely to signal by a piston displacement mechanism. The PTPSD fluctuates between piston (C-terminal helix) 'up' and piston 'down' states. Binding of an attractant to the distal PAS domain locks it in the closed form, weakening its association with the proximal domain and resulting in the transition of the latter into an open form, concomitant with a downward (towards the membrane) 4 Å piston displacement of the C-terminal helix. *In vivo*, this movement would generate a transmembrane signal by driving a downward displacement of the transmembrane helix 2 towards the cytoplasm.

1. Introduction

Most bacteria are motile. Chemotaxis, the chemically guided movement towards an attractant or away from a repellent, plays an important role in the ecology of bacterial populations. It underpins the ability of bacteria to colonize microenvironmental niches that serve as a supply of nutrients, a process central to symbiosis and pathogenesis. Chemotaxis is essential for the host colonization and virulence of many pathogenic bacteria associated with human, animal and plant diseases (Josenhans & Suerbaum, 2002; Rosenberg *et al.*, 2007). For example, chemotaxis towards chemicals released by corals and their symbionts at the endangered Great Barrier Reef is pivotal to the infection of corals by pathogenic bacteria (Rosenberg *et al.*, 2007). Renewable production of nitrogen for agriculture *via* symbiotic association between *Rhizobium* bacteria and legumes is dependent on chemotaxis towards legume roots (Fox *et al.*, 2007). Furthermore, bacterial chemotaxis plays a pivotal role in ocean-scale or global-scale biogeochemical fluxes, including carbon, nitrogen and sulfur cycling (Stocker & Seymour, 2012).



Chemical signals control the movement of bacteria *via* interaction with their membrane-embedded methyl-accepting chemotaxis protein (MCP) receptors. External chemical stimuli are detected by the periplasmic sensing domains (SDs) of the MCP receptors. Some ligands are sensed directly *via* binding to these domains and some indirectly (*e.g. via* periplasmic binding proteins), but it is the direct sensing that is of particular interest, since the specificity of the chemoreceptor ligand-binding site can, in these instances, be precisely investigated, exploited and redesigned to detect different chemical cues.

The structural basis behind the direct recognition of attractants has been systematically studied for only one family of SDs possessing a four-helix bundle fold, as exemplified by the *Salmonella typhimurium* aspartic acid receptor Tar. Previous studies have shown that binding of aspartate to the Tar SD results in a downward (towards the membrane) 1.5 Å piston displacement of the periplasmic helix. This in turn drives a piston-type sliding of the transmembrane helix 2, which links the SD to the signalling domain, towards the cytoplasm, thus transmitting the message to the cytoplasmic moiety of the receptor (Chervitz & Falke, 1996). It has since become clear that bacterial chemoreceptor SDs are extremely diverse in sequence and structure (Krell *et al.*, 2011). Two periplasmic SDs with distinctly different folds have recently been characterized: the bimodular, two four-helical bundle SD of *Pseudomonas putida* McpS (Pineda-Molina *et al.*, 2012) and the Per–Arnt–Sim (PAS)-like SD of *Helicobacter pylori* TlpB (Sweeney *et al.*, 2012). In addition, previous sequence analysis of MCPs containing the Cache (**ca**lcium channels and **che**motaxis receptors) motif (Anantharaman & Aravind, 2000) identified a new structural family of sensing modules, that of periplasmic tandem PAS SDs (PTPSDs). In the full-length receptor *in vivo*, both termini of the PTPSD are attached to transmembrane helices, one of which links the PTPSD to the cytoplasmic methyl-accepting (MA) signalling domain *via* the HAMP region (present in **h**istidine kinases, **a**denylyl cyclases, **M**CPs and **p**hosphatases). Mutagenesis studies on representative members of this family including *P. aeruginosa* PctA (Rico-Jiménez *et al.*, 2013), *Vibrio cholerae* McpX/Mlp24 (Nishiyama *et al.*, 2012), *Bacillus subtilis* McpB (Glekas *et al.*, 2010) and *Sinorhizobium meliloti* McpU (Webb *et al.*, 2014) mapped amino-acid residues in the distal PAS domain that are important for ligand recognition. However, the mechanism of signal transmission across the membrane used by SDs that differ from the well studied four-helix bundle remains to be established.

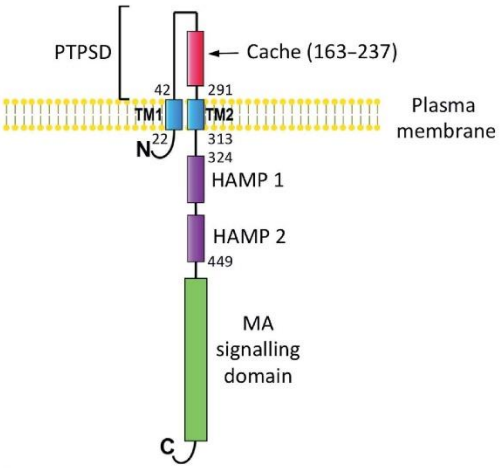


Figure 1
Overall topology of *C. jejuni* Tlp3. The PTPSD comprises residues 42–291.

Table 1
X-ray data-collection and phasing statistics.

Values in parentheses are for the outer shell.

	Gold derivative	Free form	Complex with isoleucine
Wavelength (Å)	1.04	0.95	0.95
Space group	$P2_1$	$P2_1$	$P2_1$
a, b, c (Å)	43.0, 138.2, 49.2	42.3, 137.5, 49.1	42.6, 138.0, 49.0
β (°)	94.5	94.5	94.3
Resolution range (Å)	69–1.35 (1.37–1.35)	68–1.50 (1.53–1.50)	40–1.30 (1.32–1.30)
Total No. of reflections	398772	228292	441329
No. of unique reflections	118130	75588	128260
Completeness (%)	95 (63)	85 (79)	93 (59)
Multiplicity	3.4 (2.1)	3.0 (3.1)	3.4 (2.6)
$\langle I/\sigma(I)\rangle$	9.3 (2.8)	10.5 (2.3)	12.2 (2.1)
$R_{\rm merge}$†	0.076 (0.390)	0.045 (0.281)	0.045 (0.295)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith observation of reflection hkl.

As a step towards elucidating the detailed molecular mechanism by which ligand binding to PTPSDs is signalled across the membrane, the PTPSD of the transducer-like protein 3 (Tlp3; Fig. 1) chemoreceptor from the human pathogen *Campylobacter jejuni* (Rahman *et al.*, 2014) was expressed in *Escherichia coli*, refolded from inclusion bodies, purified and crystallized as described by Machuca *et al.* (2015). This article reports the determination of the high-resolution crystal structures of the Tlp3 PTPSD and its complex with isoleucine (an attractant) and presents a model for ligand-induced conformational change within the periplasmic and membrane-spanning domains.

2. Materials and methods

2.1. Crystallization and data collection

The crystal of the complex of Tlp3 PTPSD from *C. jejuni* serotype O:2 (strain NCTC 11168) with isoleucine was obtained as described previously (Machuca *et al.*, 2015). Crystals of free Tlp3 PTPSD were produced under similar conditions. X-ray diffraction data for the native cryocooled crystal and for the isoleucine complex were collected at 100 K on the MX1 and MX2 beamlines of the Australian Synchrotron (AS; McPhillips *et al.*, 2002) to 1.5 and 1.3 Å resolution, respectively. A single-wavelength anomalous dispersion

Table 2
Refinement statistics.

	Free form	Complex with isoleucine
Resolution range (Å)	68–1.5	20–1.3
Final R_{cryst}†	0.141	0.132
Final R_{free}‡	0.189	0.165
R.m.s. deviations		
Bonds (Å)	0.015	0.019
Angles (°)	1.4	1.7
No. of atoms		
Protein	4200	4189
Water	852	800
Average B factors (Å^2)		
Protein	25	20
Water	36	34
Ligand	—	19
Ramachandran plot		
Most favoured (%)	99	97
Allowed (%)	1	3
MolProbity§ clash score	4.5	5.9
PDB code	4xmq	4xmr

† $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. ‡ The free R factor was calculated using 5% of the data omitted at random. § Chen *et al.* (2010).

(SAD) experiment was performed on a crystal of a potassium tetrabromoaurate derivative on the AS MX2 beamline. All data were processed and scaled using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (Evans & Murshudov, 2013) (see Table 1).

2.2. Structure determination

The locations of the eight Au sites for the derivative were found using *AutoSol* in *PHENIX* (Adams *et al.*, 2010). An initial partial model was generated using *AutoBuild* in *PHENIX* and was then manually completed using *Coot* (Emsley & Cowtan, 2004) and refined against the 1.3 Å resolution isoleucine data set using *PHENIX* and later *REFMAC* (Murshudov *et al.*, 2011). The structure of free Tlp3 PTPSD was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007) using the isoleucine complex as a search model. Refinement statistics and stereochemistry are given in Table 2.

2.3. Size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS)

To determine the hydrated molecular mass of Tlp3 PTPSD in solution, a 100 µl protein sample at a concentration of 100 µ*M* was loaded onto a WTC-030S5 SEC column (Wyatt Technology Corporation) pre-equilibrated with buffer 1 (100 m*M* Tris pH 8.0, 150 m*M* sodium chloride) or buffer 2 (buffer 1 plus 10 m*M* isoleucine), and the eluant was passed through an inline DAWN HELEOS light-scattering detector, an Optilab T-rEX differential refractive-index detector and a quasi-elastic light-scattering detector (WyattQELS, Wyatt Technology Corporation). For calculations of the molecular weight and R_h, the light-scattered intensity and the refractive index were analysed using *ASTRA* 6.0 (Wyatt Technology Corporation) (Table 3). Theoretical calculations of R_h from the crystal structure were carried out using *HYDROPRO* (Ortega *et al.*, 2011).

Table 3
Dynamic light-scattering results.

Sample	Polydispersity	Molecular weight (kDa)	R_h (nm)
Tlp3 PTPSD	1.0	27.0	2.5
Tlp3 PTPSD + isoleucine	1.0	27.4	2.5
BSA	1.0	63.8	3.7

2.4. Site-directed mutagenesis

Single-point alanine substitutions were introduced at positions Tyr118, Val126, Lys149, Trp151, Tyr167, Asp169, Thr170, Val171, Asp196, His237 or Arg262 *via* the oligonucleotide-directed mutagenesis technique (QuikChange, Stratagene).

2.5. CD analysis

Tlp3 PTPSD and its single-point variants were dialysed exhaustively against 10 m*M* sodium phosphate pH 7.4, 150 m*M* NaCl. Far-UV CD spectra were recorded at a protein concentration of 0.06 mg ml^{-1} at 25°C using a Jasco J-815 spectropolarimeter over the wavelength range 200–260 nm with a scan rate of 20 nm min^{-1}.

2.6. Isothermal titration calorimetry (ITC)

Tlp3 PTPSD and its variants were dialysed against buffer 1. Solutions of isoleucine (3 and 15 m*M*) were prepared in the dialysis buffer. Measurements were performed at 25°C using a VP-ITC MicroCal calorimeter (Malvern Instruments, UK). The protein sample in a 1.45 ml reaction cell was injected with 25 successive 1 µl aliquots of ligand solution at a spacing of 300 s. Binding isotherms were generated by plotting the heat change evolved per injection *versus* the molar ratio of isoleucine to Tlp3 PTPSD variant. The data were fitted to a single-site binding model using nonlinear least-squares regression, fixing the stoichiometry (N) as 1 and allowing all other fitting parameters to float (*Origin* 7, OriginLab, USA).

2.7. Bioinformatic analysis

The overall topology of Tlp3 PTPSD was obtained by searching against the SMART protein-domain database (Letunic *et al.*, 2015). To identify homologous MCP SDs, the TrEMBL database was searched for the consensus motif D*XXX*(R/K)WY*XX*A using the Quick Matrix Method in *SCANSITE* (Obenauer *et al.*, 2003) with keyword 'chemotaxis' and a molecular-weight range of 60–85 kDa. The hits were used to generate a phylogenetic tree using *PhyloT* (http://phylot.biobyte.de/) based on the NCBI taxonomy, which was visualized with the *Interactive Tree of Life* (*iTOL*; Letunic & Bork, 2011). Computational predictions of the fold of SDs were based on the detection of remote homology to proteins of known structure (*HHpred*; Söding *et al.*, 2005).

2.8. Normal-mode analysis (NMA) and morphs

NMA was performed using *elNémo* (Suhre & Sanejouand, 2004). Morphs were produced using the *UCSF Chimera* package developed by the Resource for Biocomputing,

Visualization and Informatics at the University of California, San Francisco, USA (Pettersen *et al.*, 2004).

2.9. PDB references

The atomic coordinates and structure factors for free and isoleucine-bound *C. jejuni* Tlp3 PTPSD have been deposited in the Protein Data Bank (http://www.rcsb.org; PDB entries 4xmq and 4xmr).

3. Results and discussion

3.1. Overall structure and comparison to other periplasmic SDs

The structure of *C. jejuni* Tlp3 PTPSD (residues 42–291 plus an additional N-terminal GIDPFT sequence introduced by the cloning procedure) co-crystallized with isoleucine was determined to 1.3 Å resolution using the SAD method with a gold derivative. The asymmetric unit contains a dimer with the twofold axis perpendicular to the putative membrane plane (Fig. 2*a*). The Tlp3 PTPSD subunit comprises membrane-distal and membrane-proximal PAS domains and a long stalk helix, the N-terminal and the C-terminal halves of which form part of the proximal and distal domains, respectively. Fig. 2(*b*) shows the secondary-structure topology of Tlp3 PTPSD. Although the two PAS domains do not display any significant sequence similarity, they have a very similar fold (r.m.s. deviation of 2.4 Å for the superimposition of 92 C^{α} atoms showing 10% sequence identity over equivalent positions).

The core of the distal domain (residues 63–197) comprises a central six-stranded antiparallel β-sheet with strand order 2′ 2 1 5 4 3. This β-sheet is flanked on one side by an antiparallel two-helix bundle formed by helix α2 and the C-terminal half of helix α1, and on the other side by helix α3 (secondary-structure elements are numbered to highlight deviations from the canonical PAS fold, which lacks the peripheral strand β2′). The proximal domain (residues 42–62 and 199–291) contains a central five-stranded antiparallel β-sheet with strand order 7 6 10 9 8. This β-sheet is flanked by an antiparallel two-helix bundle formed by helix α4 and the N-terminal half of helix α1 on one side and by helix α5 on the other side. In the full-length receptor, the N-terminus of helix α1 is connected to transmembrane helix 1, and the C-terminus of helix α6 is connected to transmembrane helix 2. The proximal and distal domains are intimately associated with one another, with about 17% (1170 Å^2) of the accessible surface area (ASA) of each domain buried at the interface.

This is the first crystal structure of the PTPSD of a characterized MCP receptor. It is distinctly different from the structures of the SDs of the MCPs McpS, Tar and TlpB (Fig. 3*a*). However, we note that the dimerization mode of Tlp3 PTPSD in the crystal is similar to that observed for the single PAS SD of TlpB, with the α1 and α2 helices forming a four-helix bundle at the dimerization interface. This dimerization is likely to be weak, as PTPSD is monomeric in solution according to SEC-MALS analysis (Fig. 3*b*; Table 3). Tlp3 PTPSD eluted as a single monodisperse peak both in the absence and the presence of isoleucine. The derived molecular weight was ~27 kDa, which is close to that calculated from the amino-acid sequence of a monomer (28.7 kDa). The apparent hydrodynamic radius R_h of the particles in this peak was

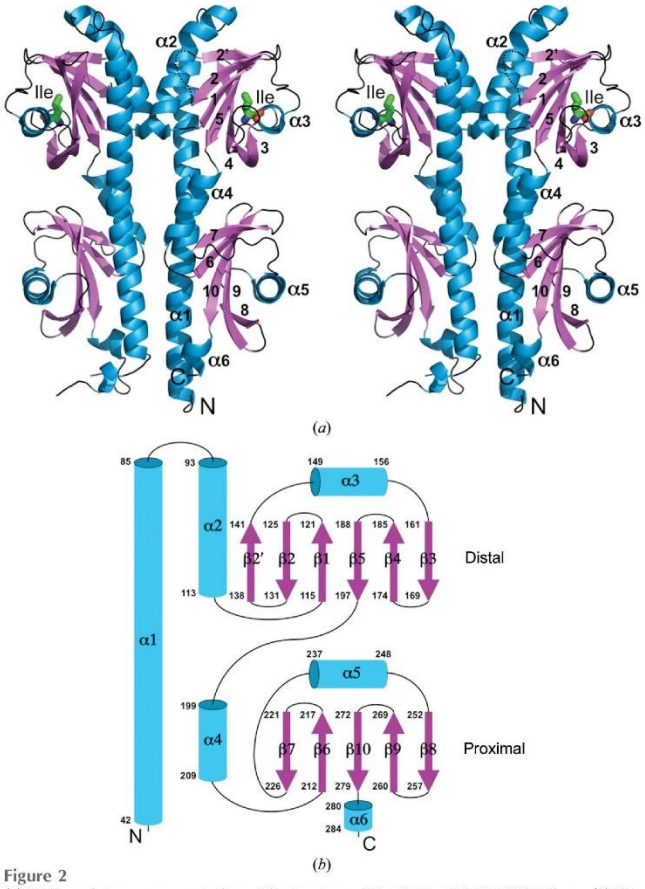


Figure 2
(*a*) Wall-eyed stereo representation of the structure of the *C. jejuni* Tlp3 PTPSD dimer. (*b*) The topology of the secondary-structure elements of Tlp3 PTPSD. The α-helices are represented by rods and β-strands by arrows. The distal and proximal PAS domains are labelled.

25 Å, which is the same as the R_h value calculated from the crystal structure of the single Tlp3 PTPSD subunit. The monomeric state of Tlp3 PTPSD in solution is in agreement with the very low relative value (6.6% or 900 Å^2) of the subunit ASA buried at the dimer interface in the crystal.

In a comparison of the atomic coordinates of Tlp3 PTPSD against the structures in the Protein Data Bank that have been characterized in the literature, using *PDBeFold* (Krissinel & Henrick, 2004), significant similarities were found with the SDs of bacterial family 1 histidine kinases (HK1$_s$s) HK1$_s$-Z2 and HK1$_s$-Z3 from *Methanosarcina mazei* (Fig. 3*c*) and HK1$_s$-Z8 from *V. parahaemolyticus* (Zhang & Hendrickson, 2010). Tlp3 PTPSD and HK1$_s$-Z2, HK1$_s$-Z3 and HK1$_s$-Z8 adopt a very similar fold [root-mean-square (r.m.s.) deviations of 2.5, 2.7 and 2.8 Å for the pairwise superimposition of 199, 195 and 200 C$^\alpha$ atoms from Z2, Z3 and Z8, respectively], despite the limited sequence homology (<18% identity for pairwise comparisons of Tlp3 PTPSD with HK1$_s$). Furthermore, a

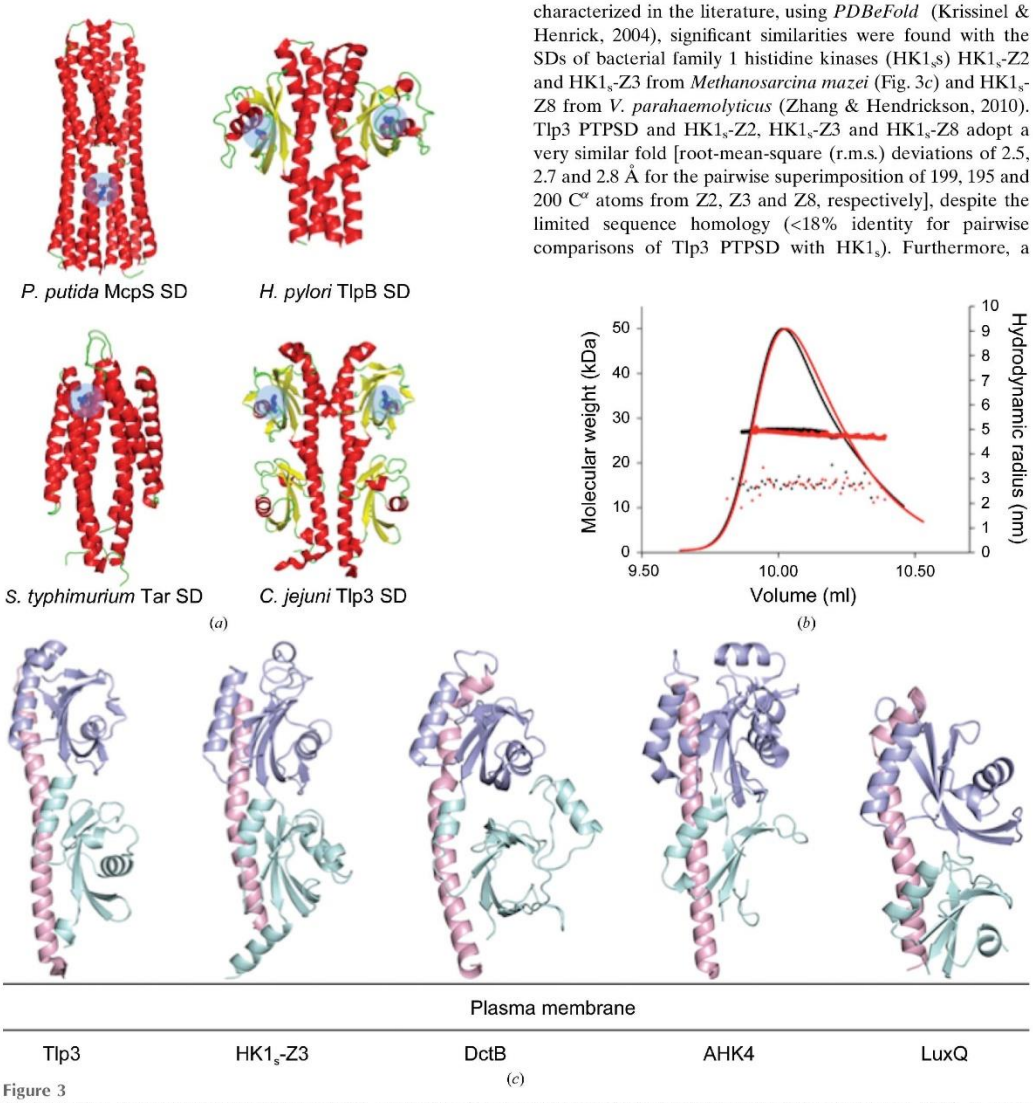


Figure 3
(*a*) Comparison of the structures of SD dimers for *P. putida* McpS (in complex with succinate; PDB entry 2yfb; Pineda-Molina *et al.*, 2012), *H. pylori* TlpB (in complex with urea; PDB entry 3ub6; Sweeney *et al.*, 2012), *S. typhimurium* Tar (in complex with aspartate; PDB entry 2lig; Milburn *et al.*, 1991) and *C. jejuni* Tlp3 (in complex with isoleucine). The bound ligands are shown in stick mode and the locations of the ligand-binding sites are highlighted with a blue circle. (*b*) Size-exclusion chromatography and molecular-weight and hydrodynamic radius (R_h) determination of Tlp3 PTPSD in the absence (black) and presence (red) of isoleucine. A bold solid line superimposed on the peak indicates the molecular weight as shown on the left-hand *y* axis. Dots represent the calculated hydrodynamic radius; its values are shown on the right-hand *y* axis. (*c*) Comparison of the structures of the periplasmic SDs of Tlp3, HK1$_s$-Z3 (PDB entry 3lib; Zhang & Hendrickson, 2010), DctB (PDB entry 3by9; Zhou *et al.*, 2008), AHK4 (PDB entry 3t4k; Hothorn *et al.*, 2011) and LuxQ (PDB entry 1zhh; Neiditch *et al.*, 2005).

similar fold comprising a long stalk helix followed by two PAS domains has previously been observed in the SDs of eukaryotic HKs, including *Arabidopsis thaliana* HK4 (AHK4; Hothorn *et al.*, 2011) and bacterial HKs from other families, including the *Sinorhizobium meliloti* C4-dicarboxylate transport sensory HK DctB (Zhou *et al.*, 2008) and the *V. harveyi* luminescence (lux) system HK LuxQ (Neiditch *et al.*, 2005), despite a very low degree of sequence conservation between those receptors and Tlp3 (Fig. 3*c*).

3.2. Isoleucine-binding site in the membrane-distal PAS domain

Analysis of the molecular surface of free Tlp3 PTPSD using *CASTp* (Dundas *et al.*, 2006) with a probe radius of 1.4 Å revealed putative ligand-binding pockets in both the membrane-distal and proximal PAS domains, with solvent-accessible volumes of 630 and 350 Å^3, respectively. The near-atomic resolution (1.3 Å) electron-density maps of the isoleucine complex of Tlp3 PTPSD are of high quality (Fig. 4*a*) and clearly show an almost identical mode of binding of isoleucine to the distal PAS domains of both subunits of the dimer in the asymmetric unit. The aliphatic side chain of isoleucine is in a largely hydrophobic environment, making van der Waals contacts with the side chains of Tyr118, Val126, Trp151 and Val171 (Fig. 4*a*) and approaching within 4.6 and 4.0 Å of the side chains of Leu128 and Leu144, respectively. The amino group forms hydrogen bonds to Tyr167 O$^{\eta}$, Asp169 O$^{\delta1}$ and Asp196 O$^{\delta2}$. The complex is further stabilized by hydrogen bonds between the carboxyl O atoms of isoleucine and Lys149 N$^{\zeta}$, Trp151 N$^{\varepsilon}$, Asp169 O$^{\delta2}$ and Thr170 O$^{\gamma}$. ASA calculations show that isoleucine is fully shielded from solvent upon binding to Tlp3.

To assess the contribution of individual amino-acid residues of Tlp3 to isoleucine binding, comprehensive alanine-scanning mutagenesis of the binding pocket was undertaken and the

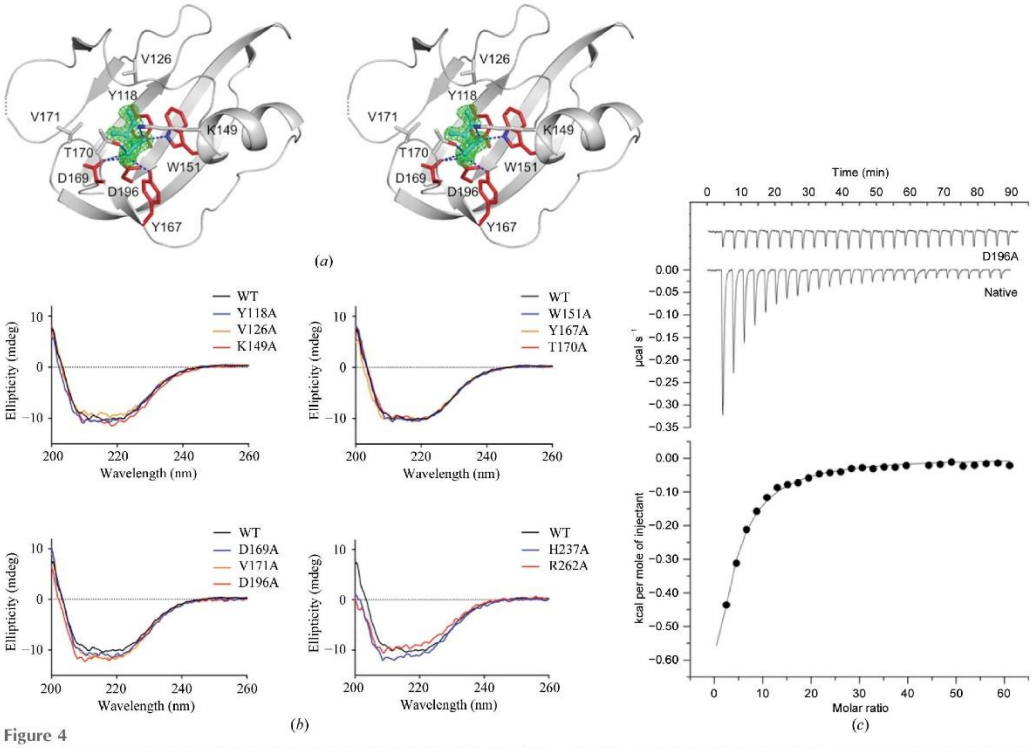


Figure 4
(*a*) The architecture of the ligand-binding site in the distal PAS domain. The $(mF_o - DF_c)\,\sigma_A$-weighted (Winn *et al.*, 2011) electron density for isoleucine is shown in green. The map was calculated at 1.3 Å resolution and contoured at the 3.0σ level. The isoleucine molecule is shown in all-atom ball-and-stick representation with C atoms coloured green. The protein side chains that form direct contacts with isoleucine are shown in stick representation. Amino-acid residues for which alanine substitutions abolished isoleucine binding are shown in red. (*b*) CD spectra of wild-type Tlp3 PTPSD (WT) and its Y118A, V126A, K149A, W151A, Y167A, D169A, T170A, V171A, D196A, H237A and R262A variants. (*c*) ITC titrations of native Tlp3 PTPSD and its D196A variant with isoleucine. Each peak in the top figure corresponds to the injection of 10 µl of 3 m*M* isoleucine into a 1.45 ml reaction cell containing protein at a concentration of 10 µ*M*. The cumulative heat of reaction is displayed in the bottom figure as a function of the ligand:protein molar ratio. The solid line is the least-squares fit of the experimental data to a single-site binding model.

effect of the substitutions on the binding affinity was examined by ITC. A total of 11 variants were generated, including two with substitutions in the proximal PAS pocket (H237A and R262A) as a negative control. The far-UV CD spectra of all of the variants were similar to that of the wild-type protein (Fig. 4*b*), indicating that they were folded and had wild-type-like secondary structure. The substitutions Y118A, W151A, Y167A, D169A and D196A in the distal PAS domain reduced the affinity of isoleucine binding by at least 35-fold (Table 4), as exemplified by the ITC titration curve for Tlp3 PTPSD D196A (Fig. 4*c*). In contrast, the substitutions V126A, K149A, T170A, V171A (distal PAS), H237A and R262A (proximal PAS) had little effect on the affinity.

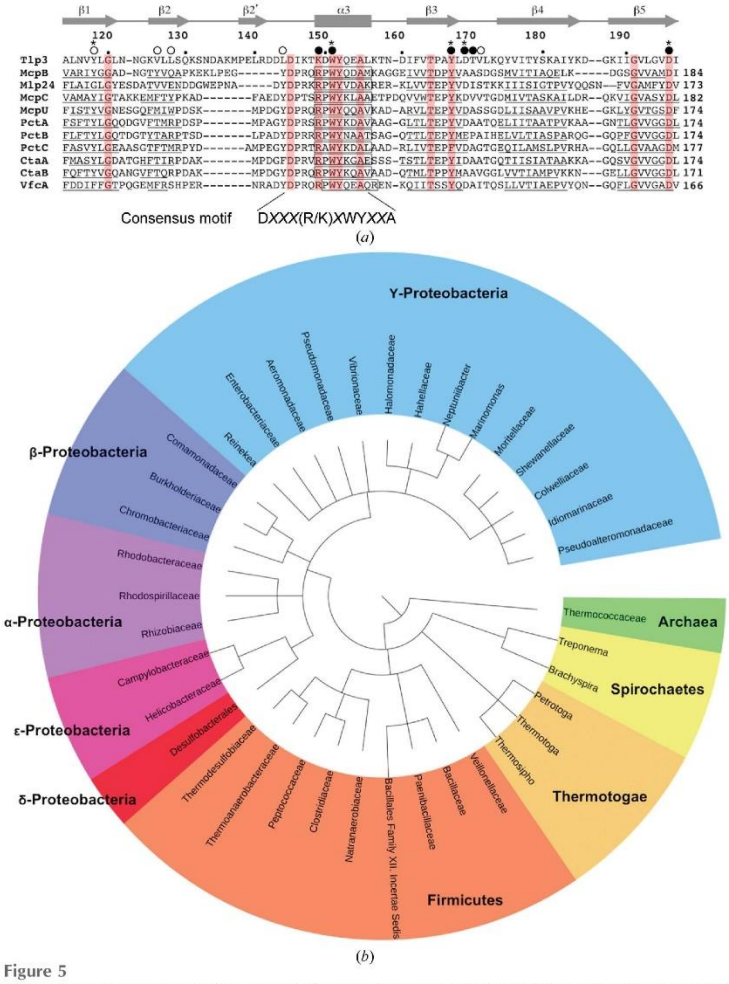


Figure 5
(*a*) Local sequence alignment of a representative subset of the characterized PTPSDs of MCP receptors for amino acids showing a consensus motif D*XXX*(R/K)*X*WY*XXA* in the ligand-binding site. The sequences are of Tlp3 from *C. jejuni*, McpB and McpC (Glekas *et al.*, 2012) from *B. subtilis*, Mlp24/McpX from *V. cholerae*, McpU from *S. meliloti*, PctA, PctB and PctC from *P. aeruginosa* (Rico-Jiménez *et al.*, 2013), CtaA and CtaB from *P. fluorescens* (Oku *et al.*, 2012) and VfcA from *V. fischeri* (Brennan *et al.*, 2013). Conserved residues are highlighted in pink. The positions of the Ala substitutions associated with a loss of isoleucine binding in Tlp3 are shown by asterisks. Residues that stabilize the invariant (main-chain) moiety and the side chain of the amino-acid ligand are marked with a filled and an open circle, respectively. (*b*) Phylogenetic distribution of species in the TrEMBL database with putative PTPSD-containing chemoreceptors for amino acids identified using the search for the consensus motif (see also Supplementary Table S1).

3.3. Comparison to other PTPSDs recognizing amino acids

To gain further insight into the structural basis of the ligand specificity of Tlp3 PTPSD, the amino-acid sequence of its ligand-binding region has been aligned with those of other characterized PTPSDs recognizing various amino acids, using their predicted secondary structures as an additional guide (Fig. 5*a*). This comparison revealed that, with the exception of Thr170, all of the protein side chains in Tlp3 PTPSD that form interactions with the amino and carboxyl groups of the ligand are strongly conserved in other amino-acid receptors containing a PTPSD. This finding suggests that the receptors of this structural family share a common mechanism of recognition of the invariant moiety of an amino acid. In contrast, the Tlp3 PTPSD residues that interact with the side chain of the ligand but do not form contacts with its invariant part are not conserved. For the known receptors with a narrow ligand specificity, there is an apparent correlation between the chemical nature of the amino acids lining the pocket for the side chain of the ligand and the physical properties of the latter. In Tlp3 PTPSD, for example, the pocket for the highly hydrophobic side chain of the isoleucine molecule is lined by the aliphatic side chains of Val126, Leu128, Leu144 and Val171. In the sequence of *B. subtilis* McpB

Table 4
Thermodynamic parameters of isoleucine binding to wild-type Tlp3 PTPSD and its variants derived from ITC measurements.

Data are means and standard deviations from two experiments.

Mutant	K_d (μM)	Enthalpy, ΔH (cal mol^{-1})	Entropy, ΔS (cal mol^{-1} K^{-1})
Native	86 ± 10	−4415 ± 230	3.6 ± 1
Y118A	>3000	ND	ND
V126A	117 ± 19	−5838 ± 148	−1.6 ± 0.8
K149A	130 ± 1	−4530 ± 105	2.6 ± 0.3
W151A	>3000	ND	ND
Y167A	>3000	ND	ND
D169A	>3000	ND	ND
T170A	51 ± 2	−6745 ± 37	−2.9 ± 0.1
V171A	129 ± 10	−9127 ± 684	−12.8 ± 2.4
D196A	>3000	ND	ND
H237A	161 ± 6	−12460 ± 42	−24.5 ± 0.2
R262A	103 ± 3	−9083 ± 338	−12.2 ± 1

(Glekas *et al.*, 2010), which is specific for asparagine and, to a lesser degree, aspartate, glutamine and histidine, these residues are substituted by those with larger, polar side chains (Tyr121, Gln123, Tyr133 and Ser161, respectively), which creates hydrogen-bonding possibilities that favour a polar amino-acid ligand smaller than isoleucine. In the sequence of the proline receptor *S. meliloti* McpU (Webb *et al.*, 2014), the side chains of Val126 and Leu128 are substituted by Met110 and Trp112, respectively, the bulky hydrophobic side chains of which are likely to protrude into the ligand-binding pocket, reducing its size while retaining its hydrophobic nature. Thus, analysis of the sequence differences between PTPSDs of Tlp3 and other MCP receptors recognizing amino acids appears to be fully consistent with the differences in their specificity.

Furthermore, analysis of the sequence alignment of a representative subset of the characterized PTPSDs of MCP receptors for amino acids identified a consensus motif in the ligand-binding site, D*XXX*(R/K)*X*WY*XX*A, that comprises helix α3 and the loop N-terminal to it. A search for this motif in the TrEMBL database, containing approximately 20 000 protein sequences, with keyword 'chemotaxis' and a molecular-weight range of 60–85 kDa identified approximately 1100 putative MCP receptors from 824 different bacteria and archaea (Fig. 5*b*; Suppementary Table S1). Analysis of the computational predictions of the membrane topology and the fold of the sensory domain confirmed that all of the hits possess a PTPSD, suggesting that the motif D*XXX*(R/K)-*X*WY*XX*A may be used as a signature for the identification of PTPSD-containing amino-acid receptors for more systematic studies.

3.4. Isoleucine-induced conformational change

Crystallization of the free protein yielded a crystal form isomorphous to that for the isoleucine complex (Table 1). The crystal structure was determined by molecular replacement using the protein monomer of the latter as a search model with the ligand and water molecules removed. Superposition of the structures of free and isoleucine-bound Tlp3 PTPSD (Fig. 6) revealed a substantial conformational change in the distal PAS domain. Its ligand-binding pocket is relatively open and accessible in the structure of the free protein. In the complex the loop connecting β3 and β4 closes over the ligand-binding site, bringing residues Tyr167, Asp169 and Thr170 into contact with the isoleucine molecule. This movement breaks a hydrogen bond and van der Waals interaction of Tyr175, residing on this loop, with the loop connecting β6 and β7, thus weakening the association between the distal and proximal domains.

3.5. Comparison of the two subunits in the asymmetric unit, normal-mode analysis and a model for the molecular mechanism of transmembrane signalling

Superposition of the structures of the two halves of the Tlp3 PTPSD dimer revealed that the proximal domain of one subunit adopts a more open form than the other, resulting in a downward ~4 Å piston displacement of the C-terminal helix α6 towards the membrane (Fig. 6). This difference in conformation was observed to the same extent in the crystals of both the free protein and its complex with isoleucine, and may be owing to different crystal-packing contacts. However, in the physiological scenario, where helices α1 at the dimer interface remain fixed perpendicular to the membrane, the different

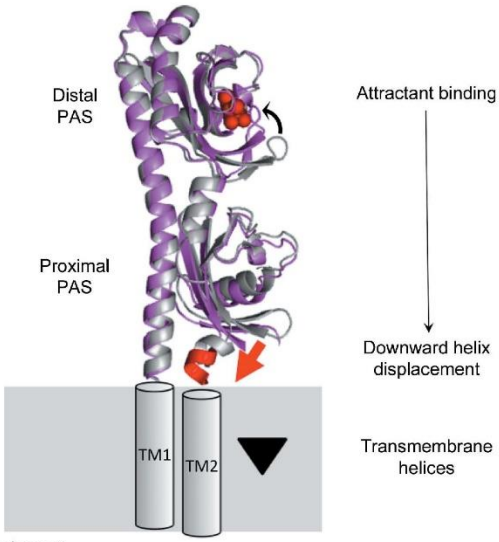


Figure 6
Piston model for transmembrane signalling by PTPSDs. The model is based on the superimposition of the two extreme conformational states of Tlp3 PTPSD observed in subunit *A* of the free protein (shown in grey) and subunit *B* of the isoleucine complex (shown in magenta/red), respectively. Attractant binding to the distal PAS domain locks it in the closed form, weakening its association with the proximal domain, which results in the transition of the latter into an open form, concomitant with a downward ~4 Å displacement of the C-terminal helix towards the membrane (see also Supplementary Movies S1 and S2). *In vivo*, this movement would be coupled to a piston-type downward displacement of the TM2 towards the cytoplasm, generating a transmembrane signal.

conformations of the proximal PAS domains provide a likely mechanism by which it can transmit the signal across the membrane *via* a downward displacement of helix α6 towards the membrane.

Analysis of the conformational changes in Tlp3 PTPSD induced by isoleucine binding suggests that the transition of the proximal domain into a more open form *in vivo* is likely to occur when its association with the distal domain is weakened by disruption of the Tyr175-mediated link. The structures of subunit *A* in the free protein (distal PAS open, proximal PAS closed, piston up) and subunit *B* in the isoleucine complex (distal PAS closed around the ligand, proximal PAS open, piston down) therefore represent the extreme conformational states of free and ligand-bound Tlp3 PTPSD in solution. To test this hypothesis, patterns of concerted movements in the Tlp3 PTPSD subunit were studied using normal-mode analysis (NMA). NMA calculations indicated that movement of one extreme conformation (free, subunit *A*) along the fourth lowest-frequency mode generated a structure similar to the other extreme conformation (isoleucine-bound, subunit *B*) (Supplementary Movies S1and S2). NMA of the movements of Tlp3 PTPSD in solution was therefore consistent with a model of signalling in which the closure of the distal PAS domain around the ligand is accompanied by the opening of the proximal domain and the displacement of its C-terminal helix towards the membrane (Fig. 6).

4. Discussion

This paper presents the first X-ray crystallographic analysis of the structure of and the ligand-induced conformational changes in the PTPSD of a characterized MCP chemoreceptor performed at near-atomic resolution (1.3 Å). Analysis of the crystal structure of the complex of *C. jejuni* Tlp3 PTPSD with the chemoattractant isoleucine reveals that upon interaction with the receptor the ligand is completely engulfed by the distal PAS domain, with extensive interactions formed between the protein and both side-chain and main-chain moieties of the ligand. Crystallographic localization of the isoleucine-binding site in the distal domain is supported by ITC experiments on the pocket mutants and is in line with previous mutagenesis studies on PTPSDs of other MCP receptors recognizing amino acids (Rico-Jiménez *et al.*, 2013; Nishiyama *et al.*, 2012; Glekas *et al.*, 2010; Webb *et al.*, 2014). Alanine substitutions of the five strongly conserved residues that stabilize the invariant (main-chain) moiety of the amino-acid ligand (Lys149, Trp151, Tyr167, Asp169 and Asp196) abrogated ligand binding both in Tlp3 PTPSD (this study) and its homologue *P. aeruginosa* PctA PTPSD (Rico-Jiménez *et al.*, 2013). This finding suggests that the mechanism of amino-acid recognition by this family of chemoreceptors is highly conserved and that the five conserved residues are important determinants that define specificity towards an amino acid, as opposed to a different small molecule.

The significance of the identification of the consensus motif in MCP receptors for amino acids is that it has now become possible to find new putative amino-acid chemoreceptors in a wide range of bacteria and archaea, which paves the way to the systematic study of this family. As one of the outcomes of this analysis, the identification of Tlp3 homologues in eubacteria (*Yersinia* spp., *Dickeya* spp.; Supplementary Table S1) dispels the notion that all enterobacterial amino-acid receptors contain four-helix bundle SDs (Rico-Jiménez *et al.*, 2013).

Our crystallographic analysis shows that although PTPSD contains a four-helix bundle at the dimer interface, like all previously characterized chemoreceptor SDs with a different fold, the bundle does not interact with the ligand. The two ligand-binding sites of the PTPSD dimer are located on the molecular faces opposite to the dimer interface. Furthermore, dimerization is not required for ligand binding, as both free and isoleucine-bound Tlp3 PTPSD are monomeric in solution, in line with previous studies on PctA PTPSD, which also binds its amino-acid ligands as a monomer (Rico-Jiménez *et al.*, 2013). Thus, the mode of ligand recognition by PTPSD is distinctly different from that of Tar, in which the ligand-binding site is located at the dimer interface (Milburn *et al.*, 1991). It is therefore particularly interesting that, despite the different fold and mode of ligand recognition, our structural analysis suggests that like the four-helix bundle SDs, PTPSDs of amino-acid chemoreceptors signal by a piston-displacement mechanism. NMA and structural comparison of the two subunits in free and isoleucine-bound Tlp3 PTPSD indicate that PTPSD may fluctuate between two conformations that correspond to the piston (C-terminal helix) 'up' and piston 'down' states. Binding of an attractant to the distal PAS domain locks it in the form closed around the ligand, which is likely to weaken its association with the proximal domain, resulting in its transition into an open form, concomitant with a downward (towards the membrane) ∼4 Å piston displacement of the C-terminal helix α6. In the full-length receptor, this movement could generate a transmembrane signal by driving a piston-type downward displacement of the transmembrane helix 2 (TM2) towards the cytoplasm.

Previous studies have shown that the two binding pockets in Tar display a negative cooperativity: the attractant binds to only one of the two, and upon binding TM2 of one monomer shifts with respect to TM2′ of the other (Chervitz & Falke, 1996). Although we have observed structural asymmetry in the two halves of the Tlp3 PTPSD dimer in the crystal, both binding pockets were occupied, possibly owing to the concentration of isoleucine in the crystallization mixture exceeding the K_d by 100-fold. It remains to be established whether binding of the attractant induces a symmetric or asymmetric change in the two halves of the Tlp3 PTPSD dimer *in vivo*. We note that earlier cross-linking studies on *B. subtilis* McpB detected no substantial change of the TM2–TM2′ interface upon asparagine binding (Szurmant *et al.*, 2004), which is consistent with a symmetric model of signalling by McpB PTPSD.

The presented structural analysis provides a foundation for more systematic mutagenesis and biochemical studies. Structural studies on the new putative amino-acid receptors of this type, as identified in this study, will be valuable. In addition, the presence of a putative 350 $Å^3$ ligand-binding pocket in the

proximal PAS domain, which opens and closes in counter-phase to the distal domain, offers a possible mechanism for modulating the receptor activity by small-molecule ligands (*e.g.* repellents).

5. Related literature

The following references are cited in the Supporting Information for this article: Alexiev *et al.* (2014), Beesley *et al.* (2010), Di Franco *et al.* (2002), Gilad (2007) and Mesbah *et al.* (2007).

Acknowledgements

We thank Victoria Korolik (Griffith University, Australia) for her contribution to the conception of the study and reagents. Part of this research was undertaken on the MX1 and MX2 beamlines of the AS, Victoria, Australia. We thank the AS staff for their assistance with data collection. This work was supported by an Australian Research Council Research Fellowship to AR (DP1094619).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Alexiev, A., Coil, D. A., Badger, J. H., Enticknap, J., Ward, N., Robb, F. T. & Eisen, J. A. (2014). *Genome Announc.* **2**, e00470-14.
Anantharaman, V. & Aravind, L. (2000). *Trends Biochem. Sci.* **25**, 535–537.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Beesley, C. A., Vanner, C. L., Helsel, L. O., Gee, J. E. & Hoffmaster, A. R. (2010). *FEMS Microbiol. Lett.* **313**, 47–53.
Brennan, C. A., DeLoney-Marino, C. R. & Mandel, M. J. (2013). *Appl. Environ. Microbiol.* **79**, 1889–1896.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Chervitz, S. A. & Falke, J. J. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 2545–2550.
Di Franco, C., Beccari, E., Santini, T., Pisaneschi, G. & Tecce, G. (2002). *BMC Microbiol.* **2**, 33.
Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acids Res.* **34**, W116–W118.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Fox, J. E., Gulledge, J., Engelhaupt, E., Burow, M. E. & McLachlan, J. A. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 10282–10287.
Gilad, J. (2007). *Recent Pat. Anti-Infect. Drug Discov.* **2**, 233–241.
Glekas, G. D., Foster, R. M., Cates, J. R., Estrella, J. A., Wawrzyniak, M. J., Rao, C. V. & Ordal, G. W. (2010). *J. Biol. Chem.* **285**, 1870–1878.
Glekas, G. D., Mulhern, B. J., Kroc, A., Duelfer, K. A., Lei, V., Rao, C. V. & Ordal, G. W. (2012). *J. Biol. Chem.* **287**, 39412–39418.
Goers Sweeney, E., Henderson, J. N., Goers, J., Wreden, C., Hicks, K. G., Foster, J. K., Parthasarathy, R., Remington, S. J. & Guillemin, K. (2012). *Structure*, **20**, 1177–1188.
Hothorn, M., Dabi, T. & Chory, J. (2011). *Nature Chem. Biol.* **7**, 766–768.
Josenhans, C. & Suerbaum, S. (2002). *Int. J. Med. Microbiol.* **291**, 605–614.
Krell, T., Lacal, J., Muñoz-Martínez, F., Reyes-Darias, J. A., Cadirci, B. H., García-Fontana, C. & Ramos, J. L. (2011). *Environ. Microbiol.* **13**, 1115–1124.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Letunic, I. & Bork, P. (2011). *Nucleic Acids Res.* **39**, W475–W478.
Letunic, I., Doerks, T. & Bork, P. (2015). *Nucleic Acids Res.* **43**, D257–D260.
Machuca, M. A., Liu, Y. C., Beckham, S. A. & Roujeinikova, A. (2015). *Acta Cryst.* F**71**, 211–216.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. (2002). *J. Synchrotron Rad.* **9**, 401–406.
Mesbah, N. M., Hedrick, D. B., Peacock, A. D., Rohde, M. & Wiegel, J. (2007). *Int. J. Syst. Evol. Microbiol.* **57**, 2507–2512.
Milburn, M. V., Privé, G. G., Milligan, D. L., Scott, W. G., Yeh, J., Jancarik, J., Koshland, D. E. Jr & Kim, S.-H. (1991). *Science*, **254**, 1342–1347.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Neiditch, M. B., Federle, M. J., Miller, S. T., Bassler, B. L. & Hughson, F. M. (2005). *Mol. Cell*, **18**, 507–518.
Nishiyama, S., Suzuki, D., Itoh, Y., Suzuki, K., Tajima, H., Hyakutake, A., Homma, M., Butler-Wu, S. M., Camilli, A. & Kawagishi, I. (2012). *Infect. Immun.* **80**, 3170–3178.
Obenauer, J. C., Cantley, L. C. & Yaffe, M. B. (2003). *Nucleic Acids Res.* **31**, 3635–3641.
Oku, S., Komatsu, A., Tajima, T., Nakashimada, Y. & Kato, J. (2012). *Microb. Environ.* **27**, 462–469.
Ortega, A., Amorós, D. & García de la Torre, J. (2011). *Biophys. J.* **101**, 892–898.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Pineda-Molina, E., Reyes-Darias, J.-A., Lacal, J., Ramos, J. L., Garcia-Ruiz, J. M., Gavira, J. A. & Krell, T. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 18926–18931.
Rahman, H., King, R. M., Shewell, L. K., Semchenko, E. A., Hartley-Tassell, L. E., Wilson, J. C., Day, C. J. & Korolik, V. (2014). *PLoS Pathog.* **10**, e1003822.
Rico-Jiménez, M., Muñoz-Martínez, F., García-Fontana, C., Fernandez, M., Morel, B., Ortega, A., Álvaro, O., Ramos, J. L. & Krell, T. (2013). *Mol. Microbiol.* **88**, 1230–1243.
Rosenberg, E., Koren, O., Reshef, L., Efrony, R. & Zilber-Rosenberg, I. (2007). *Nature Rev. Microbiol.* **5**, 355–362.
Söding, J., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
Stocker, R. & Seymour, J. R. (2012). *Microbiol. Mol. Biol. Rev.* **76**, 792–812.
Suhre, K. & Sanejouand, Y.-H. (2004). *Nucleic Acids Res.* **32**, W610–W614.
Szurmant, H., Bunn, M. W., Cho, S. H. & Ordal, G. W. (2004). *J. Mol. Biol.* **344**, 919–928.
Webb, B. A., Hildreth, S., Helm, R. F. & Scharf, B. E. (2014). *Appl. Environ. Microbiol.* **80**, 3404–3415.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zhang, Z. & Hendrickson, W. A. (2010). *J. Mol. Biol.* **400**, 335–353.
Zhou, Y.-F., Nan, B., Nan, J., Ma, Q., Panjikar, S., Liang, Y.-H., Wang, Y. & Su, X. D. (2008). *J. Mol. Biol.* **383**, 49–61.

Supporting information

Supplementary Table S1. List of species in the TrEMBL database that have one or more amino acid chemoreceptors with PTPSD.
https://doi.org/10.1107/S139900471501384X/dz5380sup1.pdf

Supplementary Movie S1. A morph movie showing the change in the structure required for the transition between the conformations observed in subunit A of the free protein and subunit B of the isoleucine complex.
https://doi.org/10.1107/S139900471501384X/dz5380sup3.mov

Supplementary Movie S2 (related to Fig. 6). NMA of free Tlp3 PTPSD (subunit A) showing movement along the fourth lowest frequency mode with amplitudes of between -30 and 100.
https://doi.org/10.1107/S139900471501384X/dz5380sup4.mov

BIOLOGICAL
CRYSTALLOGRAPHY

Volume 71 (2015)

Supporting information for article:

Structural basis for amino acid recognition and transmembrane signaling by tandem-Per-Arnt-Sim (tandem-PAS) chemoreceptor sensory domains

Yu C. Liu, Mayra P. Machuca, Simone A. Beckham, Menachem J. Gunzburg and Anna Roujeinikova

Table S1 List of species in the TrEMBL database that have one or more amino acid chemoreceptors with PTPSD.

Phylum	Class	Order	Family	Genus	Species with one or more Tlp3 homologue
			ARCHAEA		
Euryarchaeota	Thermococci	Thermococcales	Thermococcaceae	Pyrococcus	*P. horikoshii*
Euryarchaeota	Thermococci	Thermococcales	Thermococcaceae	Thermococcus	*T. kodakaraensis*
Euryarchaeota	Thermococci	Thermococcales	Thermococcaceae	Thermococcus	*T. onnurineus*
Euryarchaeota	Thermococci	Thermococcales	Thermococcaceae	Pyrococcus	*P. abyssi*
Euryarchaeota	Thermococci	Thermococcales	Thermococcaceae	Pyrococcus	*P. horikoshii*
			BACTERIA		
Spirochaetes	Spirochaetes	Spirochaetales	Spirochaetaceae	Treponema	*T. denticola*
Spirochaetes	Spirochaetes	Spirochaetales	Brachyspiraceae	Brachyspira	*B. hyodysenteriae*
Thermotogae	Thermotogales	Thermotogaceae	Petrotoga	Petrotoga	*P. mobilis*
Thermotogae	Thermotogae	Thermotogales	Thermotogaceae	Thermotoga	*T. neapolitana*
Thermotogae	Thermotogae	Thermotogales	Thermotogaceae	Thermotoga	*T. petrophila*
Thermotogae	Thermotogae	Thermotogales	Thermotogaceae	Thermosipho	*T. africanus*
Thermotogae	Thermotogae	Thermotogales	Thermotogaceae	Thermosipho	*T. melanesiensis*
Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Selenomonas	*S. flueggei*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Oceanobacillus	*O. iheyensis*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Lysinibacillus	*L. sphaericu*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. anthracis** (Beesley *et al.*, 2010)
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. clausii*

Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. cereus*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. halodurans*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. licheniformis halodurans*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. mycoides* * (Di Franko *et al.*, 2002)
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. pumilus*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. subtilis*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. thuringiensis*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	*B. weihenstephanensis*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	*A. flavithermus*
Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	*G. kaustophilus*
Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Brevibacillus	*B. brevis*
Firmicutes	Bacilli	Bacillales	Bacillales XII. Incertae Sedis	Exiguobacterium	*E. sibiricum*
Firmicutes	Clostridia	Natranaerobiales	Natranaerobiaceae	Natranaerobius	*N. thermophilus** (Mesbah *et al.*, 2007)
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium	*C. acetobutylicum*
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium	*C. beijerinckii*
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium	*C. botulinum*
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium	*C. carboxiditvorans*
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium	*C. kluyveri*
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium	*C. tetani*
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Alkaliphilus	*A. metalliredigens*
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Alkaliphilus	*A. oremlandii*
Firmicutes	Clostridia	Clostridiales	Peptococcaceae	Desulfotomaculum	*D. reducens*
Firmicutes	Clostridia	Thermoanaerobacterales	Thermoanaerobacteraceae	Thermoanaerobacter	*T. pseudethanolicus*
Firmicutes	Clostridia	Thermoanaerobacterales	Thermoanaerobacteraceae	Carboxydothermus	*C. hydrogenoformans*
Firmicutes	Clostridia	Thermoanaerobacterales	Thermoanaerobacteraceae	Caldanaerobacter	*C. subterraneus*
Firmicutes	Clostridia	Thermoanaerobacterales	Thermodesulfobiaceae	Coprothermobacter	*C. proteolyticus** (Alexiev *et al.*, 2014)

Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfovibrio	*D. magneticus*
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfovibrio	*D. salexigens*
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfovibrio	*D. vulgaris*
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfotalea	*D. psychrophila*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Helicobacter	*H. bilis*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Wolinella	*W. succinogenes*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. concisus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. curvus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. fetus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. jejuni*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. rectus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Arcobacter	*A. butzleri*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium	*A. tumefaciens*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium	*A. vitis*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	*R. meliloti*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium	*S. medicae*
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Magnetospirillum	*M. gryphiswaldense*
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Magnetospirillum	*M. magneticum*
Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Labrenzia	*L. alexandrii*
Proteobacteria	Betaproteobacteria	Neisseriales	Chromobacteriaceae	Chromobacterium	*C. violaceum*
Proteobacteria	Betaproteobacteria	Neisseriales	Chromobacteriaceae	Laribacter	*L. hongkongensis*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. ambifaria*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. cenocepacia*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. glumae*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. lata*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. mallei** (Gilad, 2007)

Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfovibrio	*D. magneticus*
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfovibrio	*D. salexigens*
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfovibrio	*D. vulgaris*
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfotalea	*D. psychrophila*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Helicobacter	*H. bilis*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Wolinella	*W. succinogenes*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. concisus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. curvus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. fetus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. jejuni*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	*C. rectus*
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Arcobacter	*A. butzleri*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium	*A. tumefaciens*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium	*A. vitis*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	*R. meliloti*
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium	*S. medicae*
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Magnetospirillum	*M. gryphiswaldense*
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Magnetospirillum	*M. magneticum*
Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Labrenzia	*L. alexandrii*
Proteobacteria	Betaproteobacteria	Neisseriales	Chromobacteriaceae	Chromobacterium	*C. violaceum*
Proteobacteria	Betaproteobacteria	Neisseriales	Chromobacteriaceae	Laribacter	*L. hongkongensis*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. ambifaria*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. cenocepacia*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. glumae*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. lata*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. mallei** (Gilad, 2007)

Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. multivorans*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. phymatum*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. phytofirmans*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. pseudomallei*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. thailandensis*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. vietnamiensis*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	*B. xenovorans*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Cupriavidus	*C. necator*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Cupriavidus	*C. pinatubonensis*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Cupriavidus	*C. taiwanensis*
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Ralstonia	*R. solanacearum*
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	*A. citrulli*
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax	*A. ebreus*
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Delftia	*D. acidovorans*
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Rhodoferax	*R. ferrireducens*
Proteobacteria	Betaproteobacteria	Burkholderiales	Ralstoniaceae	Ralstonia	*R. pickettii*
Proteobacteria	Betaproteobacteria	Burkholderiales	Ralstoniaceae	Ralstonia	*R. solanacearum*
Proteobacteria	Gammaproteobacteria	n/a	n/a	Reinekea	*R. blandensis*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Yersinia	*Y. bercovieri*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Yersinia	*Y. enterocolitica*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Yersinia	*Y. intermedia*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Yersinia	*Y. kristensenii*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Yersinia	*Y. mollaretii*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Yersinia	*Y. pseudotuberculosis*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Dickeya	*D. dadantii*
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Dickeya	*D. zeae*

Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	*S. baltica*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	*S. halifaxensis*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	*S. oneidensis*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	*S. pealeana*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	*S. putrefaciens*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	*S. sediminis*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	Shewanella	*S. woodyi*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Colwelliaceae	Colwellia	*C. psychrerythraea*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Idiomarinaceae	Idiomarina	*I. baltica*
Proteobacteria	Gammaproteobacteria	Alteromonadales	Idiomarinaceae	Idiomarina	*I. loihiensis*

*Despite the presence of motility and chemotaxis genes in their respective genomes, these species are nonmotile.

References

1 Josenhans, C. & Suerbaum, S. The role of motility as a virulence factor in bacteria. *Int J Med Microbiol* **291**, 605-614, doi:10.1078/1438-4221-00173 (2002).

2 Dasti, J. I., Tareen, A. M., Lugert, R., Zautner, A. E. & Gross, U. *Campylobacter jejuni*: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. *Int J Med Microbiol* **300**, 205-211, doi:10.1016/j.ijmm.2009.07.002 (2010).

3 Young, K. T., Davis, L. M. & Dirita, V. J. Campylobacter jejuni: molecular biology and pathogenesis. *Nat Rev Microbiol* **5**, 665-679, doi:10.1038/nrmicro1718 (2007).

4 Hugdahl, M. B., Beery, J. T. & Doyle, M. P. Chemotactic behavior of *Campylobacter jejuni*. *Infection and immunity* **56**, 1560-1566 (1988).

5 Zautner, A. E., Tareen, A. M., Gross, U. & Lugert, R. Chemotaxis in *Campylobacter jejuni*. *European journal of microbiology & immunology* **2**, 24-31, doi:10.1556/EuJMI.2.2012.1.5 (2012).

6 Zhulin, I. B. The superfamily of chemotaxis transducers: from physiology to genomics and back. *Adv Microb Physiol* **45**, 157-198 (2001).

7 Rahman, H. *et al.* Characterisation of a multi-ligand binding chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. *PLoS Pathog* **10**, e1003822, doi:10.1371/journal.ppat.1003822 (2014).

8 Li, Z. *et al.* Methyl-accepting chemotaxis proteins 3 and 4 are responsible for *Campylobacter jejuni* chemotaxis and jejuna colonization in mice in response to sodium deoxycholate. *J Med Microbiol* **63**, 343-354, doi:10.1099/jmm.0.068023-0 (2014).

9 Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic Acids Res* **42**, D222-230, doi:10.1093/nar/gkt1223 (2014).

10 Anantharaman, V. & Aravind, L. Cache - a signaling domain common to animal Ca(2+)-channel subunits and a class of prokaryotic chemotaxis receptors. *Trends Biochem Sci* **25**, 535-537 (2000).

11 Glekas, G. D. *et al.* A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. *J. Biol. Chem.* **285**, 1870-1878, doi:10.1074/jbc.M109.072108 (2010).

12 Nishiyama, S. *et al.* Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infection and immunity* **80**, 3170-3178, doi:10.1128/IAI.00039-12 (2012).

13 Zhang, Z. & Hendrickson, W. A. Structural characterization of the predominant family of histidine kinase sensor domains. *Journal of molecular biology* **400**, 335-353, doi:10.1016/j.jmb.2010.04.049 (2010).

14 Machuca, M. A., Liu, Y. C., Beckham, S. A. & Roujeinikova, A. Cloning, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for multiple ligands (CcmL). *Acta crystallographica. Section F, Structural biology communications* **71**, 211-216, doi:10.1107/S2053230X1500045X (2015).

15 Liu, Y. C., Machuca, M. A., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. *Acta crystallographica. Section D, Biological crystallography* **71**, 2127-2136, doi:10.1107/S139900471501384X (2015).

Chapter 3: The crystal structure of the dCache ligand binding domain of *C. jejuni* transducer-like protein 1 (Tlp1)

Preface for Chapter 3

Chemotactic movement of *C. jejuni* towards conditions that favour its growth has been shown to be directed by mucine components (L-fucose and L-serine), tricarboxylic acid (TCA) cycle intermediates (citrate, fumarate, α-ketoglutarate and succinate pyruvate) and several amino acids (aspartate, glutamate, cysteine and isoleucine) [1,2]. Most of these molecules are metabolised by *C. jejuni* [3,4]. However, amino acids are known as the major carbon sources used by this bacterium, since most of *C. jejuni* strains are unable to metabolise exogenous sugars [5,6].

Biochemical studies have shown that *C. jejuni* preferentially utilises L-aspartic acid and L-serine [7] as carbon and/or energy sources, which can be deaminated into intermediates that enters the TCA cycle [8]. *C. jejuni* exhibits chemotactic response towards aspartate, which is mediated by Tlp1 receptor [also termed as CcaA (*Campylobacter* chemoreceptor for aspartate A) or Cj1506c][9]. Tlp1 is ubiquitous and the most conserved chemoreceptor among *C. jejuni* strains [10,11]. However, the LBD of Tlp1 shares no sequence similarity with that of the *E. coli* aspartate receptor Tar or any other chemoreceptor described to date. Therefore, the mechanism involved in the aspartate sensing by Tlp1 may be unique to this bacterial genus. The following series of experiments were designed to: elucidate the crystal structure of Tlp1; and identify the mechanism for ligand recognition by this chemoreceptor.

This chapter comprises a crystallographic report published in December 2014 in the *Acta Crystallographica Section F* that describes the cloning, overexpression, refolding, purification, crystallisation, preliminary analysis of the X-ray diffraction data [12]. The detailed structural, biophysical and biochemical analysis is presented in the journal article format and was published in February 2016 in the *Journal of Structural Biology* [13].

Monash University

Declaration for Thesis Chapter 3 (Section 3.1)

Declaration by candidate

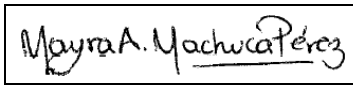
In the case of Chapter 3 (section 3.1), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiments, performed experiments, analysed data and wrote manuscript	53

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Yu C. Liu	Designed experiments, analysed data and wrote manuscript	
Anna Roujeinikova	Designed experiments, analysed data, wrote manuscript and led the research	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's Signature		**Date** 10 /12/17
Main Supervisor's Signature		**Date** 13/12/17

Section 3.1.

Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

Cloning, expression, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for aspartate A (CcaA)

Mayra A. Machuca,[a] Yu C. Liu[a] and Anna Roujeinikova[a,b]*

[a]Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia, and [b]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

Correspondence e-mail: anna.roujeinikova@monash.edu

Received 15 November 2014
Accepted 15 December 2014

In *Campylobacter jejuni*, chemotaxis and motility have been identified as important virulence factors that are required for host colonization and invasion. Chemotactic recognition of extracellular signals is mediated by the periplasmic sensory domains of its transducer-like proteins (Tlps). In this study, the sensory domain of the *C. jejuni* chemoreceptor for aspartate A (CcaA) has been expressed in *Escherichia coli* and purified from inclusion bodies. The urea-denatured protein was refolded and then crystallized by the hanging-drop vapour-diffusion method using PEG 3350 as a precipitating agent. A complete data set has been collected to 1.4 Å resolution using cryocooling conditions and synchrotron radiation. The crystals belonged to space group *P*1, with unit-cell parameters a = 39.3, b = 43.3, c = 50.9 Å, α = 92.5, β = 111.4, γ = 114.7°.

1. Introduction

Campylobacter jejuni is a leading cause of bacterial foodborne enterocolitis in humans (Zilbauer *et al.*, 2008). *C. jejuni* colonizes many wild and domestic animals, and is also found in untreated water (Boes *et al.*, 2005; Oporto *et al.*, 2007). In humans, infection occurs after the ingestion of contaminated food, especially poultry products (Hepworth *et al.*, 2011). Symptoms include fever, abdominal pain and diarrhoea (Acheson & Allos, 2001). Notably, *Campylobacter* enteritis has been associated with several post-infection complications, including Guillain–Barré syndrome, reactive arthritis and idiopathic peripheral neuropathy (Schmidt-Ott *et al.*, 2006; Friedman *et al.*, 2000).

The molecular mechanisms of *C. jejuni* pathogenesis are still poorly understood. However, flagellar motility and chemotaxis are known to play an important role during the colonization and infection of both avian and mammalian hosts (Josenhans & Suerbaum, 2002; Dasti *et al.*, 2010). In chemotaxis, swimming of motile bacteria may be controlled by various extracellular chemical gradients. These external stimuli are detected by the sensory domains of methyl-accepting chemotactic proteins (MCPs) or transducer-like proteins (Tlps), which transfer information through their signalling domains, activating signalling cascades that control the direction of rotation of flagella (Zhulin, 2001; Fernando *et al.*, 2007; Zautner *et al.*, 2012).

The *C. jejuni* genome encodes at least ten different Tlps, which have been classified into three groups according to their domain organization: A (Tlp1, Tlp2, Tlp3, Tlp4, $Tlp7_{mc}$, $Tlp7_{m}$ and Tlp10), B (Tlp9) and C (Tlp5, Tlp6, $Tlp7_{c}$ and Tlp8) (Parkhill *et al.*, 2000; Zautner *et al.*, 2012). Sequence-homology analysis has shown that the signalling domain is highly conserved in Tlps from different bacterial species; in contrast, the sensory domain, which is involved in the recognition of external ligands, is highly varied (Marchant *et al.*, 2002; Parkhill *et al.*, 2000), which reflects evolution to detect a broad spectrum of environmental cues.

To date, the ligand specificity of four *C. jejuni* Tlps has been characterized. Tlp1 has been identified as an aspartate receptor and termed the *Campylobacter* chemoreceptor for aspartate A (CcaA; Hartley-Tassell *et al.*, 2010), Tlp3 as a multiligand-binding receptor (Rahman *et al.*, 2014; Li *et al.*, 2014), Tlp4 as a sodium deoxycholate receptor and Tlp7 as a formic acid receptor (Tareen *et al.*, 2010). The periplasmic sensory domain of Tlp1 (CcaA) has no homology to that of Tar, the well characterized aspartate receptor of *Escherichia coli*

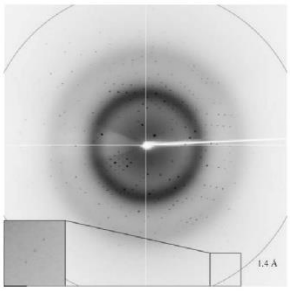


Table 1
Macromolecule-production information.

Source organism	*C. jejuni* serotype O:2 (strain NCTC 11168)
DNA source	Plasmid pGU0605 (Hartley-Tassell *et al.*, 2010)
Forward primer	CACCACAAAACAAGTAAGTCAAAATATTACC
Reverse primer	CTCGAGTTATTTTTTTAAGGGTTTAAATACTGAATA
Cloning vector	pET151/D-TOPO
Expression vector	pET151/D-TOPO
Expression host	*E. coli* strain BL21-CodonPlus (DE3)-RIPL
Complete amino-acid sequence of the construct produced	GIDPFTKQVSQNITKNTEDILASITKEYATQTQGIFGEMIAL-NKSISGTLTEMFRSTSKEDLDIDNITNIITNTFDNSAYSN-FTYLYLIDPPEYFKEESKFFNTQSGKFVMLYADEEKDNKG-GIKAIQASDEIANLQVVQDILKKAKYGENKVYIGRPIKMN-LEGQDFDAVNVAIPIFDRKNQVVGVIGMTLDFSDIATYLL-DPKGQKYDGELRVLLNSDGFMAIHPNKNLVLKNLKDINPN-KGAQETYKAISEGKNGVFNYIASDGDDSYAAINSFKVQDS-SWAVLVTAPKYSVFKPLKK

(Hartley-Tassell *et al.*, 2010), or any of the chemoreceptors described to date. Therefore, the mechanism involved in aspartate recognition by *C. jejuni* CcaA may be unique to this bacterial genus. In order to understand the molecular basis of ligand recognition by this novel sensory domain, we have expressed, refolded from inclusion bodies, purified and crystallized the periplasmic domain of CcaA (CcaA^{peri}).

2. Materials and methods

2.1. Gene cloning, protein expression, refolding and purification

CcaA^{peri} (amino-acid residues 31–327; UniProtKB Q0P8B2; also termed Tlp1 or cj1506c) was cloned and overexpressed in *E. coli*. Briefly, the DNA fragment encoding CcaA^{peri} was amplified from the pGU0605 plasmid (Hartley-Tassell *et al.*, 2010) using the primers tlp1-F (5′-CACCCACAAAACAAGTAAGTCAAAATATTACC-3′) and tlp1-R (5′-CTCGAGTTATTTTTTTAAGGGTTTAAATACTG-AATA-3′) (Table 1). The PCR product was ligated into the pET151/D-TOPO vector using the TOPO cloning kit (Invitrogen) to generate an expression plasmid that contained an N-terminal His_6 tag followed by a TEV protease cleavage site. The correct insertion of the fragment was verified by DNA sequencing.

E. coli BL21(DE3)-RIPL cells (Stratagene) were transformed with the expression vector and cultured to exponential phase at 37°C in LB medium containing 50 µg ml^{-1} ampicillin. Expression was induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside at an $OD_{600\ nm}$ of 0.6. The cells were grown for a further 4 h at 37°C and then harvested by centrifugation at 6000*g* for 15 min at 4°C. The cell pellet was resuspended in 100 ml buffer *A* (10 m*M* Tris–HCl pH 8.0) and lysed using an Avestin EmulsiFlex-C5 high-pressure homogenizer. The cell lysate was centrifuged at 10 000*g* for 15 min at 4°C. Analysis of the clarified supernatant and pellet by SDS–PAGE showed that the recombinant CcaA^{peri} was predominantly deposited in inclusion bodies. We have recently reported a procedure for the production of a soluble, crystallizable form of the periplasmic sensory domain of the bacterial chemoreceptor TlpC by extraction from inclusion bodies and refolding (Liu & Roujeinikova, 2014). To purify CcaA^{peri} for crystallization, we used a similar protocol with minor variations. The inclusion-bodies pellet was washed three times with buffer *B* [10 m*M* Tris–HCl pH 8.0, 0.2 m*M* phenylmethanesulfonyl fluoride (PMSF), 1% Triton X-100] and once with buffer *C* (10 m*M* Tris–HCl pH 8.0, 0.2 m*M* PMSF). The inclusion bodies were then solubilized in buffer *D* [10 m*M* Tris–HCl pH 8.0, 8 *M* urea, 10 m*M* dithiothreitol (DTT), 0.2 m*M* PMSF] for 30 min at 4°C with mixing by axial rotation. The mixture was then clarified by centrifugation at 30 000*g* for 30 min at 4°C and the concentration of solubilized protein was determined using the Bradford assay (Bradford, 1976).

CcaA^{peri} refolding was performed by diluting 60 mg denatured protein (typical volume 1–2 ml) into 250 ml buffer *E* (3 *M* urea, 100 m*M* Tris–HCl pH 8.0, 0.4 *M* L-arginine monohydrochloride) followed by 48 h incubation at 4°C with continuous mixing. The protein sample was then dialyzed against 5 l buffer *A* for 16 h at 4°C. NaCl, Tris–HCl pH 8.0 and imidazole were added to the protein solution to final concentrations of 500, 20 and 20 m*M*, respectively, and the sample was loaded onto a 5 ml HiTrap Chelating HP column (GE Healthcare) equilibrated with buffer *F* (500 m*M* NaCl, 20 m*M* Tris–HCl pH 8.0, 20 m*M* imidazole). The column was washed with 30 column volumes of the same buffer and the protein was eluted with buffer *G* (500 m*M* NaCl, 20 m*M* Tris–HCl pH 8.0, 500 m*M* imidazole).

The N-terminal hexahistidine tag was cleaved off with His_6-TEV protease (Invitrogen) whilst dialyzing the sample overnight at 4°C against buffer *H* [150 m*M* NaCl, 10 m*M* Tris–HCl pH 8.0, 2 m*M* DTT, 1%(*v*/*v*) glycerol]. NaCl and imidazole were then added to the sample to final concentrations of 500 and 20 m*M*, respectively, and the TEV protease and the uncleaved protein were removed on a HiTrap Chelating HP column. The flowthrough fractions containing protein were pooled, concentrated to 400 µl in an AmiconUltra Ultracel 10 kDa cutoff concentrator and passed through a Superdex 200 HiLoad 26/60 gel-filtration column (GE Healthcare) equilibrated with buffer *I* (10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl) at a flow rate of 0.4 ml min^{-1}. The protein purity was estimated by SDS–PAGE to be greater than 95% (Fig. 1).

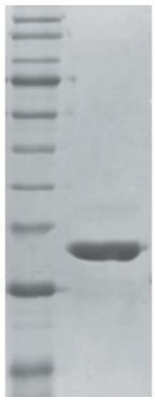


Figure 1
Reduced 12% SDS–PAGE Coomassie Blue-stained gel of purified recombinant CcaA^{peri}. The left lane contains molecular-weight marker (labelled in kDa).

2.2. Crystallization

Prior to crystallization, the protein was concentrated to 10 mg ml^{-1}, mixed with 10 m*M* aspartic acid and centrifuged for 20 min at 13 000*g* to clarify the solution. Crystallization screening was performed by the hanging-drop vapour-diffusion method using an automated Phoenix crystallization robot (Art Robbins Instruments) and JBScreen HTS I and II (Jena Bioscience), The JCSG+ Suite (Qiagen), Crystal Screen HT and PEG/Ion HT (Hampton Research). 200 nl protein solution was mixed with 200 nl reservoir solution and equilibrated against 50 µl reservoir solution in a 96-well Art Robbins CrystalMation Intelli-Plate (Hampton Research). Crystals appeared in many different conditions. The best-looking crystals grew in

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	96-well Art Robbins CrystalMation Intelli-Plate (Hampton Research)
Temperature (K)	293
Protein concentration (mg ml^{-1})	10
Buffer composition of protein solution	100 m*M* ammonium acetate, 12%(*w*/*v*) PEG 3350, 50 m*M* bis-tris pH 5.5
Composition of reservoir solution	200 m*M* ammonium acetate, 23%(*w*/*v*) PEG 3350, 100 m*M* bis-tris pH 5.5

200 m*M* ammonium acetate, 25%(*w*/*v*) PEG 3350, 100 m*M* bis-tris pH 5.5 (The JCSG+ Suite condition H10). Upon refinement of this condition to improve the quality of the crystals, the concentration of PEG was reduced to 23%(*w*/*v*); crystals appeared in the manually set up trays (drop size 2 µl) after two weeks. Crystallization information is summarized in Table 2.

2.3. Data collection and processing

The crystals were briefly soaked in a cryoprotectant solution consisting of 220 m*M* ammonium acetate, 26%(*w*/*v*) PEG 3350, 100 m*M* bis-tris pH 5.5, 10 m*M* aspartic acid, 10%(*v*/*v*) glycerol and were flash-cooled by plunging them into liquid nitrogen. X-ray diffraction data were collected to 1.4 Å resolution on the MX1 beamline of the Australian Synchrotron (Fig. 2). A total of 360 images were collected using 0.5° oscillations. The data were processed and scaled using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (Evans & Murshudov, 2013) from the *CCP*4 suite (Winn *et al.*, 2011). Calculation of the self-rotation function was performed using *POLARRFN* (Winn *et al.*, 2011). The statistics of data collection are summarized in Table 3.

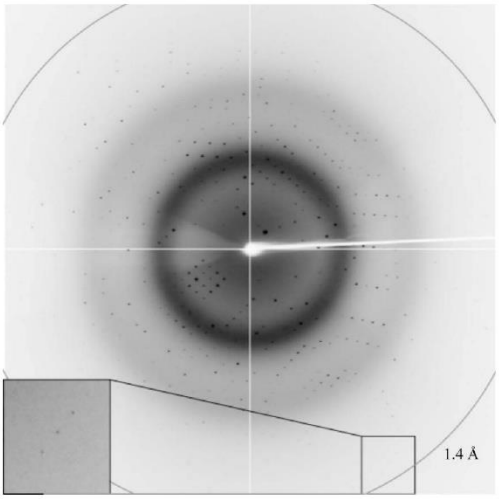


Figure 2
A representative 0.5° oscillation image obtained from a CcaA^{peri} crystal using a CCD ADSC Quantum 315r detector on beamline MX1 of the Australian Synchrotron. The magnified rectangle shows diffraction spots at 1.4 Å resolution.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	MX1 beamline, Australian Synchrotron
Wavelength (Å)	1.0
Temperature (K)	100
Detector	ADSC Quantum 315r CCD
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	1
Space group	*P*1
a, *b*, *c* (Å)	39.3, 43.3, 50.9
α, β, γ (°)	92.5, 111.4, 114.7
Mosaicity (°)	0.6
Resolution range (Å)	38.3–1.4 (1.42–1.40)
Total No. of reflections	97980 (3448)
No. of unique reflections	49677 (1741)
Completeness (%)	92 (66)†
Multiplicity	2 (2)
$\langle I/\sigma(I)\rangle$	11.9 (3.9)
$R_{r.i.m.}$	0.046 (0.150)
Overall *B* factor from Wilson plot (Å^2)	8.6

† The low completeness in *P*1 resulted from the use of a single-axis goniometer.

3. Results and discussion

Recombinant CcaA^{peri} from *C. jejuni* was overexpressed in *E. coli* BL21(DE3)-RIPL cells. The purified protein fragment contained CcaA residues 31–327 plus six additional residues at the N-terminus (a cloning artifact; GIDPFT). When subjected to gel filtration, the protein eluted as a single peak with a retention volume corresponding to an approximate molecular weight of 30 kDa, which is close to the theoretical molecular weight (33.6 kDa), suggesting that recombinant CcaA^{peri} is a monomer in solution under the tested conditions.

Crystals of CcaA^{peri} were obtained in the presence of its ligand (aspartic acid) using a sparse-matrix crystallization approach. Data collected from a cryocooled crystal at the Australian Synchrotron showed diffraction to 1.4 Å resolution (Fig. 2). Autoindexing of the diffraction data using *iMosflm* showed that it belonged to space group *P*1, with unit-cell parameters a = 39.3, b = 43.3, c = 50.9 Å, α = 92.5, β = 111.4, γ = 114.7°. The average $I/\sigma(I)$ value was 11.9 for all reflections (resolution range 38.3–1.4 Å) and 3.9 in the highest resolution shell (1.42–1.40 Å). A total of 50 732 measurements were made of 49 677 independent reflections. Data processing gave an $R_{r.i.m.}$ of 0.046 for intensities (0.15 in the 1.42–1.40 Å resolution shell) and these data were 92% complete (66% completeness in the highest resolution shell).

Calculations of the Matthews coefficient (Matthews, 1977) and solvent content, assuming the presence of one molecule in the unit cell, gave values of 2.2 Å^3 Da^{-1} and 43%, respectively. Analysis of the self-rotation function calculated using data in the resolution range 10–5 Å with an integration radius of 25 Å revealed no dominant features that could be confidently assigned to noncrystallographic axes. Together, this analysis suggests that the CcaA^{peri} crystals contain one molecule per unit cell. Structure determination by molecular replacement was not possible, as a sequence-similarity search against the Protein Data Bank did not identify any homologues of known structure. A search for heavy-atom derivatives suitable for structure determination by the multiple isomorphous replacement and/or multiwavelength anomalous dispersion methods is under way.

Plasmid pGU0605 was kindly provided by Professor Victoria Korolik (Griffith University, Australia). We thank the staff at the Australian Synchrotron for their assistance with data collection. We also thank Dr Danuta Maksel and Dr Robyn Gray at the Monash

Crystallography Unit for assistance in setting up robotic crystallization trials. AR is an Australian Research Council Research Fellow.

References

Acheson, D. & Allos, B. M. (2001). *Clin. Infect. Dis.* **32**, 1201–1206.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Boes, J., Nersting, L., Nielsen, E. M., Kranker, S., Enøe, C., Wachmann, H. C. & Baggesen, D. L. (2005). *J. Food Prot.* **68**, 722–727.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Dasti, J. I., Tareen, A. M., Lugert, R., Zautner, A. E. & Gross, U. (2010). *Int. J. Med. Microbiol.* **300**, 205–211.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Fernando, U., Biswas, D., Allan, B., Willson, P. & Potter, A. A. (2007). *Int. J. Food Microbiol.* **118**, 194–200.
Friedman, C., Neimann, J., Wegener, H. & Tauxe, R. (2000). *Campylobacter*, edited by I. Nachamkin & M. J. Blaser, pp. 121–138. Washington: ASM Press.
Hartley-Tassell, L. E., Shewell, L. K., Day, C. J., Wilson, J. C., Sandhu, R., Ketley, J. M. & Korolik, V. (2010). *Mol. Microbiol.* **75**, 710–730.
Hepworth, P. J., Ashelford, K. E., Hinds, J., Gould, K. A., Witney, A. A., Williams, N. J., Leatherbarrow, H., French, N. P., Birtles, R. J., Mendonca, C., Dorrell, N., Wren, B. W., Wigley, P., Hall, N. & Winstanley, C. (2011). *Environ. Microbiol.* **13**, 1549–1560.
Josenhans, C. & Suerbaum, S. (2002). *Int. J. Med. Microbiol.* **291**, 605–614.
Li, Z., Lou, H., Ojcius, D. M., Sun, A., Sun, D., Zhao, J., Lin, X. & Yan, J. (2014). *J. Med. Microbiol.* **63**, 343–354.
Liu, Y. C. & Roujeinikova, A. R. (2014). *Protein Expr. Purif.* **107**, 29–34.
Marchant, J., Wren, B. & Ketley, J. (2002). *Trends Microbiol.* **10**, 155–159.
Matthews, B. W. (1977). *X-ray Structure of Proteins*, 3rd ed., edited by H. Neurath & R. L. Hill, Vol. 3, pp. 468–477. New York: Academic Press.
Oporto, B., Esteban, J. I., Aduriz, G., Juste, R. A. & Hurtado, A. (2007). *J. Appl. Microbiol.* **103**, 977–984.
Parkhill, J. *et al.* (2000). *Nature (London)*, **403**, 665–668.
Rahman, H., King, R. M., Shewell, L. K., Semchenko, E. A., Hartley-Tassell, L. E., Wilson, J. C., Day, C. J. & Korolik, V. (2014). *PLoS Pathog.* **10**, e1003822.
Schmidt-Ott, R., Schmidt, H., Feldmann, S., Brass, F., Krone, B. & Gross, U. (2006). *Clin. Vaccine Immunol.* **13**, 779–783.
Tareen, A. M., Dasti, J. I., Zautner, A. E., Gross, U. & Lugert, R. (2010). *Microbiology*, **156**, 3123–3135.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zautner, A. E., Tareen, A. M., Gross, U. & Lugert, R. (2012). *Eur. J. Microbiol. Immunol.* **2**, 24–31.
Zhulin, I. B. (2001). *Adv. Microb. Physiol.* **45**, 157–198.
Zilbauer, M., Dorrell, N., Wren, B. W. & Bajaj-Elliott, M. (2008). *Trans. R. Soc. Trop. Med. Hyg.* **102**, 123–129.

Monash University

Declaration for Thesis Chapter 3 (Section 3.2)

Declaration by candidate

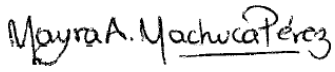
In the case of Chapter 3 (section 3.2), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiments, performed experiments, analysed data and wrote manuscript	53

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Yu C. Liu	Designed experiments, analysed data and wrote manuscript	
Simone A. Beckham	Analysed data, provided intellectual input and reviewed manuscript	
Menachem J. Gunzburg	Analysed data, provided intellectual input and reviewed manuscript	
Anna Roujeinikova	Designed experiments, analysed data, wrote manuscript and led the research	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's Signature		**Date** 10/12/17
Main Supervisor's Signature		**Date** 13 /12/17

Section 3.2.

Journal of Structural Biology 194 (2016) 205–213

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

The crystal structure of the tandem-PAS sensing domain of *Campylobacter jejuni* chemoreceptor Tlp1 suggests indirect mechanism of ligand recognition

Mayra A. Machuca [a,b], Yu C. Liu [a,b], Simone A. Beckham [a,c], Menachem J. Gunzburg [a,c], Anna Roujeinikova [a,b,c,*]

[a] Infection and Immunity Program, Monash Biomedicine Discovery Institute, Australia
[b] Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia
[c] Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

ARTICLE INFO

Article history:
Received 19 November 2015
Received in revised form 22 February 2016
Accepted 23 February 2016
Available online 24 February 2016

Keywords:
Chemotaxis
Methyl-accepting protein
Sensing domain
Crystal structure

ABSTRACT

Chemotaxis and motility play an important role in the colonisation of avian and human hosts by *Campylobacter jejuni*. Chemotactic recognition of extracellular signals is mediated by the periplasmic sensing domain of methyl-accepting chemotactic proteins (membrane-embedded receptors). In this work, we report a high-resolution structure of the periplasmic sensing domain of transducer-like protein 1 (Tlp1), an aspartate receptor of *C. jejuni*. Crystallographic analysis revealed that it contains two Per-Arnt-Sim (PAS) subdomains. An acetate and chloride ions (both from the crystallisation buffer) were observed bound to the membrane-proximal and membrane-distal PAS subdomains, respectively. Surprisingly, despite being crystallised in the presence of aspartate, the structure did not show any electron density corresponding to this amino acid. Furthermore, no binding between the sensing domain of Tlp1 and aspartate was detected by microcalorimetric experiments. These structural and biophysical data suggest that Tlp1 does not sense aspartate directly; instead, ligand recognition is likely to occur indirectly *via* an as yet unidentified periplasmic binding protein.

1. Introduction

Campylobacter jejuni is a Gram-negative, microaerophilic, flagellated bacterium that colonises the intestines of many wild and domestic animals (Boes et al., 2005; Oporto et al., 2007). *C. jejuni* is a leading cause of bacterial foodborne gastroenteritis in humans (Zilbauer et al., 2008). Human infection occurs by the consumption of contaminated food (especially poultry products) or water (Hepworth et al., 2011). Although this infection is usually self-limiting, it may lead to important post-infection complications such as neuromuscular paralysis (Guillain–Barré syndrome), reactive arthritis, myositis and idiopathic peripheral neuropathy (Friedman et al., 2000; Schmidt-Ott et al., 2006).

The molecular mechanisms of *C. jejuni* pathogenesis are still poorly understood. However, flagellar motility and chemotaxis are known to play a crucial role during the intestinal colonisation of both avian and mammalian hosts (Dasti et al., 2010; Josenhans and Suerbaum, 2002). Non-motile strains of *C. jejuni* could not, for example, colonise the digestive tract of mice (Morooka et al., 1985). Similarly, chemotactic defects such as the loss of chemoreceptor function resulted in reduced motility and infectivity of *C. jejuni* (Hendrixson et al., 2001; Tareen et al., 2010).

Chemotactic movement of *C. jejuni* towards conditions that favour survival has been shown to be directed by various extracellular chemical gradients. These include L-fucose and L-serine (mucin components), citrate, fumarate, α-ketoglutarate and succinate (tricarboxylic acid (TCA) cycle intermediates) and amino acids, such as aspartate and glutamate, that may be deaminated to TCA cycle intermediates (Hugdahl et al., 1988; Sebald and Vernon, 1986). These external stimuli are detected by the sensing domains of methyl-accepting chemotactic proteins (MCPs), or transducer-like proteins (Tlps), which transfer information through their signaling domains, activating signaling cascades that

Abbreviations: MCP, methyl-accepting chemotactic protein; Tlp, transducer-like protein; PAS, Per-Arnt-Sim; PTPSD, periplasmic tandem PAS sensing domain; ITC, isothermal titration calorimetry; CD, circular dichroism.

* Corresponding author at: 19 Innovation Walk, Monash University, Clayton, Victoria 3800, Australia.
E-mail address: Anna.Roujeinikova@monash.edu (A. Roujeinikova).

http://dx.doi.org/10.1016/j.jsb.2016.02.019
1047-8477/

control the direction of rotation of flagella (Fernando et al., 2007; Zautner et al., 2012; Zhulin, 2001).

The *C. jejuni* genome encodes at least 11 different putative Tlps, which have been classified into three groups according to their domain organisation: A (Tlp1, Tlp2, Tlp3, Tlp4, Tlp7$_{mc}$, Tlp7$_{m}$, Tlp10 and Tlp11); B (Tlp9); and C (Tlp5, Tlp6, Tlp7$_{c}$ and Tlp8) (Parkhill et al., 2000; Zautner et al., 2012). The group B chemoreceptor is a single-span membrane protein, with no periplasmic domain, whereas group C chemoreceptors are soluble cytoplasmic proteins. As such, group B and C receptors likely sense internal signals such as, for example, changes in redox potential. Group A chemoreceptors have a general membrane topology reminiscent of that of the well-characterised chemoreceptors of *Escherichia coli*. They are predicted to comprise an N-terminal transmembrane helix followed by a periplasmic domain (putative ligand-binding, or sensing domain), a second transmembrane helix and, finally, a C-terminal cytoplasmic signaling domain, which contains a methyl-accepting chemotaxis-like subdomain (chemotaxis sensory transducer) (Fig. 1). The presence of a periplasmic domain in group A Tlps suggests that they sense external signals. The cytoplasmic signaling domain is highly conserved in Tlps from different bacterial species (Marchant et al., 2002; Parkhill et al., 2000). In contrast, the periplasmic sensing domains of *C. jejuni* share no significant sequence similarity with those of *E. coli*, which suggests that the mechanisms by which they recognise ligands are likely distinctly different from *E. coli*.

The periplasmic sensing domains, involved in the recognition of external ligands, are highly diverse across different bacterial species, which reflects evolution to detect a broad spectrum of environmental cues (Wadhams and Armitage, 2004). To date, some of the ligands of five *C. jejuni* Tlps from this group have been identified. Tlp1 has been shown to be involved in sensing aspartate and termed the *Campylobacter* chemoreceptor for aspartate A (CcaA) (Hartley-Tassell et al., 2010). Tlp3 has been shown to directly sense isoleucine and may indirectly recognise several other small molecules (Li et al., 2014; Rahman et al., 2014). Tlp4 has been identified as a sodium deoxycholate receptor, Tlp7 as a formic acid receptor (Tareen et al., 2010), and Tlp11 as a galactose receptor (Day et al., 2014).

Tlp1 is ubiquitous and the most conserved among *C. jejuni* strains (Day et al., 2014; Korolik and Ketley, 2008). Previous studies showed that the *tlp1* gene was strongly upregulated in the *C. jejuni* strains passaged through chicken (King et al., 2013), and its insertional inactivation resulted in significantly decreased colonisation of the chicken intestinal tract (Hartley-Tassell et al., 2010). Furthermore, the *tlp1* mutant strain of *C. jejuni* showed drastically reduced ability to invade human epithelial cells and chicken embryo (Vegge et al., 2009). Tlp1 is therefore involved in the commensal colonisation of the chicken intestine by *C. jejuni* and may play an important role in the infection and colonisation of the human host.

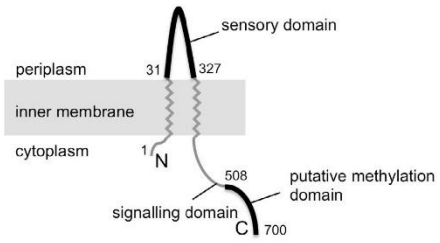


Fig. 1. The predicted membrane topology of *C. jejuni* Tlp1 and the boundaries (amino acid residue numbers) of the periplasmic sensing domain and the putative methylation domain.

The *tlp1*- isogenic mutant of *C. jejuni* showed a significant decrease in chemotaxis towards aspartate (Hartley-Tassell et al., 2010). Migration towards aspartate could be restored to this mutant by complementation with an intact copy of the *tlp1* gene. These findings suggested that Tlp1 plays an important role in sensing aspartate. The periplasmic sensing domain of Tlp1 has no sequence homology with that of Tar, the well-characterised aspartate receptor of *E. coli* (Hartley-Tassell et al., 2010), or any of the chemoreceptors described to date. Therefore, the mechanism involved in the aspartate recognition by *C. jejuni* Tlp1 may be unique to this bacterial genus.

Recently, we determined the crystal structure of the sensing domain of *C. jejuni* Tlp3 and its complex with isoleucine (Liu et al., 2015), and analysis of its ligand-binding pocket provided a molecular explanation for the specificity of that chemoreceptor. Here, we report the 1.4-Å resolution crystal structure of the periplasmic region of *C. jejuni* Tlp1 comprising amino acid residues 31–327. The structure revealed that, like Tlp3, Tlp1 contains a periplasmic tandem-Per-Arnt-Sim (PAS) sensing domain (PTPSD). Our structural and biophysical analysis suggests that, in contrast to Tlp3, the mechanism of amino acid recognition by Tlp1 does not involve direct ligand-receptor interaction; aspartate is likely sensed by Tlp1 indirectly, *via* an as yet unidentified periplasmic binding protein.

2. Materials and methods

2.1. Crystallisation and data collection

Tlp1-PTPSD (residues 31–327 plus GIDPFT sequence at the N-terminus as a cloning artifact) was crystallised in the presence of 10 mM L-aspartic acid as described (Machuca et al., 2015). The crystals belong to space group *P*1, with unit cell parameters $a = 39.3$ Å, $b = 43.3$ Å, $c = 50.9$ Å, $\alpha = 92.5°$, $\beta = 111.4°$, $\gamma = 114.7°$ and a monomer in the asymmetric unit. Although we were unable to crystallise Tlp1-PTPSD in the absence of aspartate under similar conditions, we noted that addition of 10 mM aspartic acid to the crystallisation buffer reduced pH from 5.5 to 5.2. To test if it was lower pH, rather than the presence of aspartic acid, that was important for the crystal formation, we prepared a crystallisation mix that had the same components (minus aspartic acid), except the pH of the buffer was 5.2 (23% PEG 3350, 200 mM ammonium acetate and 100 mM Bis-Tris pH 5.2). Tlp1-PTPSD did crystallise under these conditions, and the crystals had similar morphology to those grown in the presence of aspartic acid (unit cell parameters $a = 39.0$ Å, $b = 43.7$ Å, $c = 50.4$ Å, $\alpha = 92.7°$, $\beta = 110.6°$, $\gamma = 115.0°$). Crystals obtained in the presence of aspartate (the putative Tlp1 ligand) were selected for the subsequent analysis.

To perform data collection at cryogenic temperatures, the crystals were briefly soaked in a cryoprotectant solution consisting of 220 mM ammonium acetate, 26% (*w*/*v*) PEG 3350, 100 mM Bis-Tris pH 5.5, 10 mM aspartic acid and 10% (*v*/*v*) glycerol, and flash-cooled by plunging into liquid nitrogen. X-ray diffraction data were collected to 1.4 Å using the MX1 beamline of the Australian Synchrotron (AS). The diffraction data were processed and scaled using *iMOSFLM* (Battye et al., 2011) and AIMLESS (Evans and Murshudov, 2013) from the CCP4 software suite (Winn et al., 2011). Data collection statistics are summarised in Table 1.

2.2. Structure determination

The structure of Tlp1-PTPSD was determined using PHASER (McCoy et al., 2005) with coordinates of *Helicobacter pylori*

Table 1
X-ray data collection and processing statistics. Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.0
Resolution (Å)	1.4 (1.42 – 1.40)
Total no. of observed reflections	97,980 (3448)
No. of unique reflections	49677 (1741)
Completeness (%)	92 (66)
Multiplicity	2.0 (2.0)
Mean $I/\sigma(I)$	11.9 (3.9)
R_{merge} [a]	0.046 (0.147)
Overall *B* factor from Wilson plot (Å^2)	8.6

[a] $R_{merge} = \frac{(\sum_h \sum_i |I_{hi} - \langle I_h \rangle|)}{\sum_h \sum_i I_{hi}|}$, where I_{hi} is the intensity of the *i*th observation of reflection *h*.

TlpC-PTPSD (Y.C.L., M.A.M. and A.R., unpublished data) as a search model. The model comprising residues 37–142, 148–326 was built through iterative cycles of rebuilding with COOT (Emsley and Cowtan, 2004) and refinement with PHENIX (Adams et al., 2010). Subsequent difference Fourier maps clearly revealed density corresponding to two ligands other than water molecules. The density in the membrane-proximal and membrane-distal domains was interpreted as an acetate and chloride ion, respectively (both being crystallisation buffer components). The average B-factors for the acetate and chloride ion in the final refined model (10.5 and 9.3 Å^2, respectively) are similar to that of the surrounding protein, suggesting that both ions are bound at close to full occupancy. Analysis of the stereochemical quality of the model was performed using MOLPROBITY (Chen et al., 2010). The final refined model of Tlp1-PTPSD contains 285 of the expected 303 amino acid residues,

Table 2
Refinement statistics.

Resolution range (Å)	15 – 1.4
Reflections	47,127
Atoms	2878
Water molecules	414
[a]*R*-factor	0.117
[b]R_{free}	0.154
Bond-length deviation from ideality (Å)	0.014
Bond-angle deviation from ideality (°)	1.3
Molprobity scores	
Ramachandran regions (%)	
Favoured	99
Allowed	1
Outliers	0
Clashscore	4.3
Average B (protein atoms) (Å^2)	15.6
Average B (water molecules) (Å^2)	32.9
Average B (acetate ion) (Å^2)	10.5
Average B (chloride ion) (Å^2)	9.3

[a] $R = \frac{\sum_h ||F_{obs}| - |F_{calc}||}{\sum_h |F_{obs}|}$.

[b] The free *R*-factor was calculated on 5% of the data omitted at random.

one acetate and one chloride ion and 414 water molecules (*R*-factor 0.117, R_{free} 0.154). All the non-glycine residues lie in permitted regions of the Ramachandran plot, with 99% of these in the most favoured regions. Refinement statistics are given in Table 2. Structure figures were prepared using PYMOL (Schrödinger, 2016).

2.3. Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis

The absolute hydrated molecular mass of Tlp1-PTPSD in solution was determined by MALS analysis coupled to size-exclusion chromatography. Tlp1-PTPSD was dialysed against buffer containing 100 mM sodium acetate pH 5.5 and 200 mM ammonium acetate, and concentrated to 2.9 mg ml^{-1}. A 100 µl sample was loaded onto a WTC-030S5 size-exclusion column (Wyatt Technology Corporation) pre-equilibrated with the same buffer flowing at 0.4 ml min^{-1}. The eluate was passed through an in-line DAWN HELEOS light scattering detector, an Optilab T-rEX differential refractive index detector and a quasi-elastic light scattering detector (WyattQELS, Wyatt Technology Corporation). The experiment was repeated in the presence of aspartic acid, which was added to the concentration of 10 mM to both dialysis and size-exclusion chromatography buffers. A bovine serum albumin (BSA) standard was used to normalise the MALS detectors. For calculations of the molecular mass and hydrodynamic radius, the light-scattered intensity and refractive index were analysed using ASTRA 6.0 (Wyatt Technology Corporation), with a value for the refractive index increment (dn/dc) protein of 0.185 ml g^{-1}. The results are presented in Table 3. Theoretical calculations of the hydrodynamic radius from the crystal structure were carried out using HYDROPRO version 10 (Ortega et al., 2011).

2.4. Circular dichroism spectroscopy

Purified Tlp1-PTPSD was buffer-exchanged into either 10 mM sodium phosphate (pH 7.4) or 10 mM sodium phosphate (pH 7.4) plus 150 mM NaCl using dialysis. Far-UV circular dichroism (CD) spectra were recorded at a protein concentration of 0.05 mg/ml at 25 °C using a JASCO J-815 spectropolarimeter over a wavelength range 200–260 nm with a scan rate of 20 nm/min. Spectra were recorded in triplicate and averaged. The data were corrected for solvent and smoothed using the Savitzky–Golay algorithm with a radius of 25.

2.5. Isothermal titration calorimetry (ITC)

Purified Tlp1-PTPSD was extensively dialysed against either buffer 1 (10 mM Tris/HCl pH 8.0) for the experiment with chloride or buffer 2 (10 mM Tris/HCl pH 8.0, 150 mM NaCl) for the experiments with acetate and sodium salt of L-aspartic acid. 150 mM solution of NaCl for ITC measurements was prepared in the dialysis buffer 1 after the dialysis was complete. 8 mM solution of ammonium acetate and 30 mM solution of sodium salt of L-aspartic acid were prepared in the dialysis buffer 2. Measurements were

Table 3
Molecular weights and hydrodynamic radii determined by SEC MALLS/QELS analysis.

	Tlp1-PTPSD	Tlp1-PTPSD + aspartic acid
Buffer composition	Ammonium acetate 200 mM Sodium acetate 100 mM pH 5.5	Ammonium acetate 200 mM Sodium acetate 100 mM pH 5.5 Aspartic acid 10 mM
Theoretical MW (kDa)	33.6	33.6
Experimental MW (kDa)	31.4	31.7
Theoretical hydrodynamic radius (nm)	2.8	2.8
Experimental hydrodynamic radius (nm)	2.6	2.6
Polydispersity	1.0	1.0

performed at 25 °C using a VP-ITC MicroCal calorimeter (Malvern Instruments, UK). The protein sample in a 1.45-ml reaction cell (concentration 110 μM) was injected with 25 successive 10-μl aliquots of the respective ligand at a spacing of 300 s. Based on our analysis of the crystal structure of Tlp1-PTPSD, the number of binding sites for the chloride and acetate ions was presumed to be 1. Binding isotherms were generated by plotting the heat change evolved per injection *versus* molar ratio of ligand to Tlp1-PTPSD. The data was fitted to a single-site binding model using non-linear least-squares regression, fixing the stoichiometry (N) as 1 and floating all other fitting parameters (Origin 7, OriginLab, USA).

2.6. Protein Data Bank accession number

The coordinates of Tlp1-PTPSD have been deposited in the RCSB Protein Data Bank under accession code 4WY9.

3. Results and discussion

3.1. Overall fold of Tlp1-PTPSD

Tlp1-PTPSD crystals diffracted to 1.4-Å resolution at the MX1 station of the Australian Synchrotron. The Tlp1-PTPSD subunit contains a long stalk helix followed by a repeat of two PAS domains (Fig. 2). The N-terminal and C-terminal halves of the stalk helix form part of the membrane-proximal and membrane-distal PAS domains, respectively. The overall fold of Tlp1-PTPSD is presented in Fig. 2(A and B). Fig. 2C represents the topology of the secondary structure elements of Tlp1-PTPSD deduced from the φ–ψ values calculated with the program PROCHECK (Laskowski et al., 1993).

The two PAS domains have a similar, although not identical, fold. The distal domain (residues 60–218) contains a central six-stranded antiparallel β-sheet, with the strand order 2′ 2 1 5 4 3. This β-sheet is flanked on one side by helix α2′ and an antiparallel two-helix bundle formed by helix α2 and the C-terminal half of helix α1, and on the other side, by helices α3 and α3′ (secondary structure elements are numbered to highlight deviations from the canonical PAS fold, which lacks the peripheral strand β2′ and the helical insertion between strands β1 and β2). The core of the proximal domain (residues 31–59, 219–327) comprises a central five-stranded antiparallel β-sheet, with the strand order 7 6 10 9 8. This β-sheet is flanked by an antiparallel two-helix bundle formed by helix α4 and the N-terminal half of helix α1 on one side, and helices α5 and α5′ on the other side, whereas the shorter helix α6 forms a C-terminal extension of strand β10. *In vivo*, the N-terminus of helix α1 is connected to the Tlp1 transmembrane helix 1, and the C-terminus of helix α6 to transmembrane helix 2. The membrane-proximal and membrane-distal domains are intimately associated with each other by jointly contributing apolar residues to the large hydrophobic cluster at their interface (Fig. 2D).

3.2. Tlp1-PTPSD is monomeric in solution and the crystal

Crystals of Tlp1-PTPSD belonged to space group *P*1 and contained one molecule in the unit cell, indicating that Tlp1-PTPSD is monomeric in the crystal. To determine the oligomeric state of Tlp1-PTPSD in solution, multiangle light scattering (MALS) analysis coupled to size-exclusion chromatography was carried out in buffer with the salt concentration and pH identical to those in the crystallisation mix. Both in the presence of 10 mM aspartate and without it, Tlp1-PTPSD eluted as a single, close to symmetrical, peak (Fig. 3) with a polydispersity index value of 1, indicating that the eluted particles were homogenous with respect to the molar mass. The elution volumes with and without aspartate were almost identical. The derived MW value for Tlp1-PTPSD in solution was 31.4 ± 2.5 kDa (Table 3; we estimated the accuracy of the weight determination as approx. 8% from the quality of the BSA standard). This value is consistent with a monomer. Calculation of the Stokes radius from the crystal structure of the Tlp1-PTPSD monomer gives the value of 2.8 nm, which is very close to the value derived from the MALS analysis (2.6 nm both with and without aspartate). These results demonstrate that *C. jejuni* Tlp1-PTPSD exists as a monomer in solution, which is in line with the previous reports of the monomeric behaviour of PTPSDs of *C. jejuni* Tlp3 (Liu et al., 2015) and *Pseudomonas putida* PctA (Rico-Jimenez et al., 2013).

3.3. Comparison to other periplasmic sensing domains

A similar tandem arrangement of two PAS domains has been previously observed in periplasmic sensing domains (SDs) of chemoreceptors and histidine kinases, despite a very low degree of sequence conservation between those receptors and Tlp1. In comparison of Tlp1-PTPSD against the published structures deposited in the Protein Data Bank, using Fold (Krissinel and Henrick, 2004), the closest structural similarities were found with SDs of bacterial family 1 histidine kinases ($HK1_S$) $HK1_S$-Z2, $HK1_S$-Z3 from *Methanosarcina mazei* (Fig. 4) and $HK1_S$-Z8 from *Vibrio parahaemolyticus* (Zhang and Hendrickson, 2010). Tlp1-PTPSD and $HK1_S$-Z2, $HK1_S$-Z3 and $HK1_S$-Z8 adopt a similar fold (root mean square (rms) deviation of 2.8 Å, 2.7 Å and 2.9 Å for the pairwise superimposition of 232, 242 and 223 C_α atoms from Z2, Z3 and Z8, respectively), while showing less than 17% identity over equivalenced positions.

Further comparisons show that a similar fold comprising a long stalk helix followed by two PAS domains has been previously observed with SDs of bacterial HKs from other families, including: *Sinorhizobium meliloti* C4-**dic**arboxylate **t**ransport sensory HK DctB (Zhou et al., 2008) and *Vibrio harveyi* luminescence (lux) system HK LuxQ (Neiditch et al., 2005); different bacterial chemoreceptors, including *C. jejuni* Tlp3 (Liu et al., 2015); and eukaryotic HKs, including *Arabidopsis thaliana* HK 4 (AHK4) (Hothorn et al., 2011) (Fig. 4). These proteins appear to be remote functional homologues derived from different ancestral genes that have undergone evolutionary convergence to similar folds.

3.4. Analysis of small-molecule ligands bound to Tlp1

Two non-protein moieties were identified in the Tlp1-PTPSD crystal structure, both presumably derived from the crystallisation solution. A single ion (shown as a green sphere in Figs. 2A and 5A) has been identified in the structure of the membrane-distal PAS domain as a strong peak on the difference map. The ion-binding site is located in a small pocket formed by L161, Q162, V163, N197, I184 and G181, with the solvent-accessible volume of 82 $Å^3$ (calculated using CASTp (Dundas et al., 2006)). The ion is stabilised by hydrogen bonds to the main-chain amide of N163 and three ordered water molecules (Fig. 5A). The lack of protein electron negative groups in the vicinity suggested that this is unlikely to be a positively charged metal ion. When it was modelled as Cl^-, its refined *B* factor (9.3 $Å^2$) was close to the average *B* factor of surrounding protein atoms (9.8 $Å^2$). Chloride binding to Tlp1-PTPSD was confirmed by ITC (Fig. 5B). When the binding isotherm for the interaction was fitted by nonlinear regression using a simple bimolecular interaction model, the data were consistent with a 1:1 interaction between the chloride ion and protein with a dissociation constant of ~70 mM. As this value lies within the range of chloride concentration in the host gut environment, we have addressed the question about the role of the chloride ion in the structure. The membrane-distal PAS domain in other SDs has previously been implicated in direct interaction with

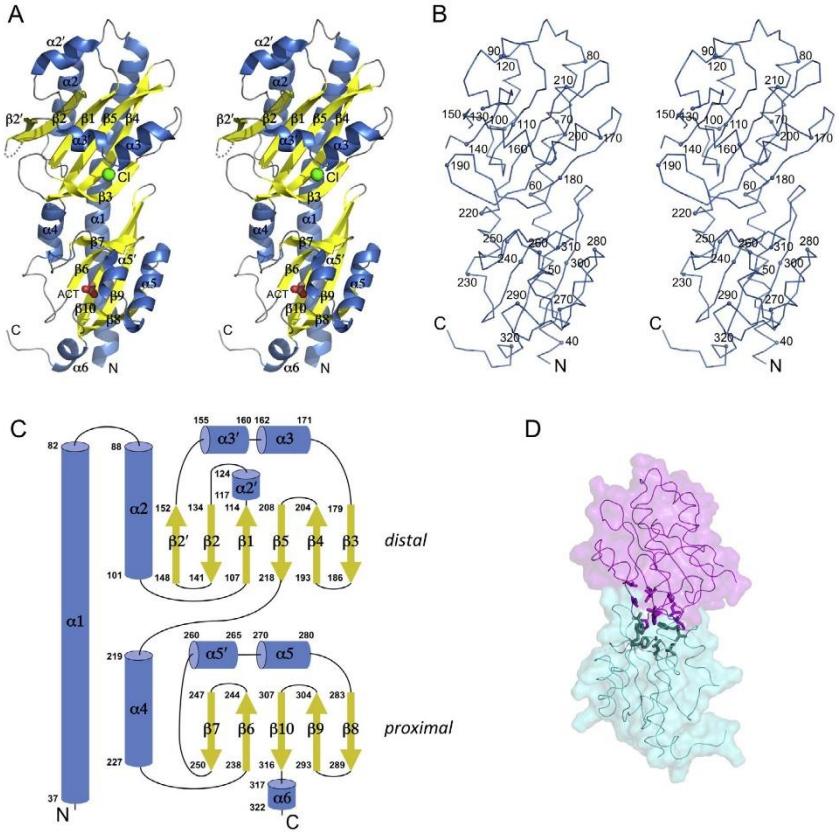


Fig. 2. The overall fold of *C. jejuni* Tlp1-PTPSD. (A) Stereo representation of Tlp1-PTPSD. β-Strands and α-helices are represented as arrows and coils. The chloride ion is represented as a green ball and the acetate ion (ACT) is shown with the carbon atoms coloured orange, and oxygen atoms red. (B) C_α backbone trace of Tlp1-PTPSD with every tenth residue labelled. (C) The topology of secondary structure elements of Tlp1-PTPSD. The α-helices are represented by rods and β-strands by arrows. (D) The hydrophobic cluster formed by residues I60, F61, M64, I65, V178, I180, P183, V196, V198, L217 from the membrane-distal domain and F219, I222, L241, A249, I250, F302, V304, W309 from the membrane-proximal domain.

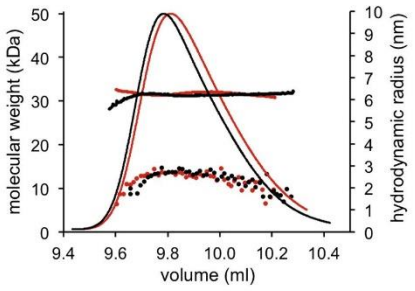


Fig. 3. Size-exclusion chromatography and molecular weight (MW) and hydrodynamic (Stokes) radius determination of Tlp1-PTPSD in the absence (black) and presence (red) of aspartic acid. A solid line superimposed on the peak indicates the MW as shown on the left-hand *y*-axis. Dots represent the hydrodynamic radius R_h calculated over the central portion of the elution peak (shown by UV trace). The R_h values are shown on the right-hand *y*-axis.

small-molecule chemoattractants (Glekas et al., 2010; Liu et al., 2015; Nishiyama et al., 2012; Rico-Jimenez et al., 2013; Webb et al., 2014). However, it has been reported that chloride does not elicit chemotactic response in *C. jejuni* (Hugdahl et al., 1988). We note that the chloride ion links the helix α3 with the central β-sheet (Fig. 2A), and that Tlp1-PTPSD showed less secondary structure in the absence of chloride than in the presence of 150 mM chloride as judged by circular dichroism (Fig. 5C). Together with the fact that the amino-acid residues that form the chloride-binding pocket are conserved in different strains of *C. jejuni* (Supplementary Fig. 1), this suggests that chloride is probably needed to maintain the conformation of Tlp1-PTPSD.

An acetate ion (from the crystallisation solution) has been identified in the structure of the membrane-proximal PAS domain (Fig. 6A) at a site that is similar to the ligand-binding pocket of other PAS domains (Henry and Crosson, 2011). The methyl group of the acetate is bound in a hydrophobic pocket, making van der Waals contacts with the side chains of L238, V240, L261 and L312. The hydroxyl group of acetate is in a hydrophilic environment and forms hydrogen bonds with the side chains of H251,

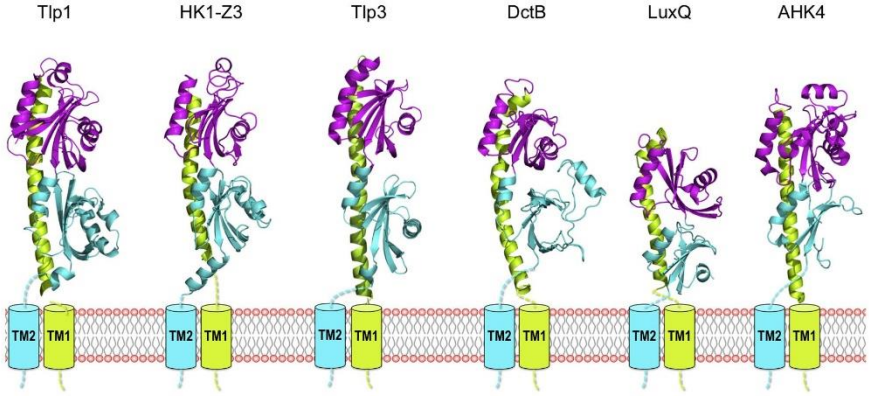


Fig. 4. Comparison of the structures of the periplasmic SDs of Tlp1, HK1$_{\chi}$-Z3 (PDB code 3LIB, (Zhang and Hendrickson, 2010)), Tlp3 (PDB code 4XMQ, (Liu et al., 2015)), DctB (PDB code 3BY9, (Zhou et al., 2008)), LuxQ (PDB code 1ZHH, (Neiditch et al., 2005)), and AHK4 (PDB code 3T4K, (Hothorn et al., 2011)).

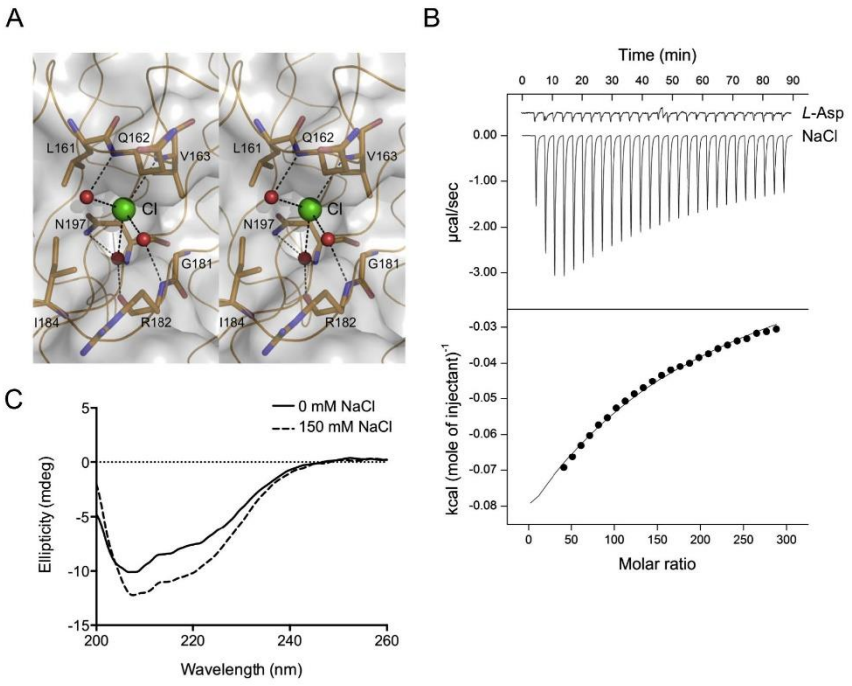


Fig. 5. Chloride binding to Tlp1-PTPSD. (A) Stereoview of the chloride-binding site in the membrane-distal PAS domain. (B) ITC analysis of aspartate (top trace) and Cl^- (bottom trace) binding to Tlp1-PTPSD. Each peak corresponds to the injection of 10 µl of either 30 mM solution of sodium salt of L-aspartic acid or 150 mM NaCl into a 1.45-ml reaction cell containing protein at a concentration of 110 µM. Cumulative heat of reaction is displayed in the bottom figure as a function of the ligand/protein molar ratio. The solid line is the least-squares fit of the experimental data to a single-site binding model. (C) CD spectra (ellipticity) for Tlp1-PTPSD in the buffer containing either 10 mM sodium phosphate pH 7.4 (solid line) or 10 mM sodium phosphate pH 7.4 and 150 mM NaCl (dashed line). Each CD spectrum is the average of three replicate scans.

Y287 and S290, and with the main-chain amide of S290. We have confirmed acetate binding to Tlp1-PTPSD by ITC (Fig. 6B). The data could be fitted to a single-site binding model that yielded the K_D value of 3.4 mM. However, acetate does not elicit chemotactic response in *C. jejuni* (Hugdahl et al., 1988), and therefore it is unlikely to be a natural signal for Tlp1. The solvent-accessible volume of the pocket occupied by the acetate (88 Å^3) is small compared to, for example, the isoleucine-binding pocket of the PAS domain of

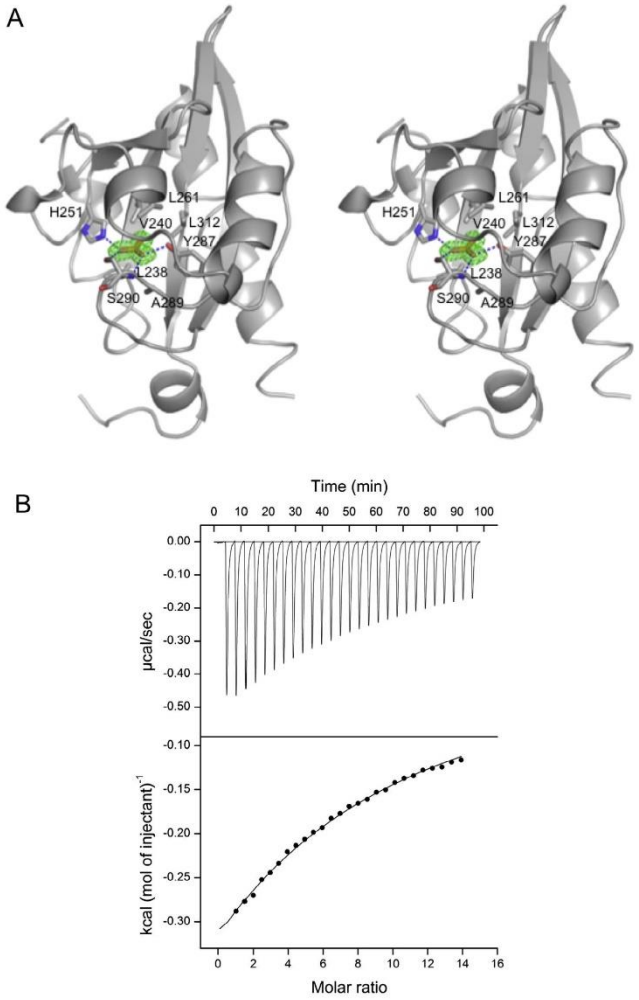


Fig. 6. Acetate binding to Tlp1-PTPSD. (A) The ($mF_o - DF_c$) σ_A-weighted (Winn et al., 2011) electron density for the acetate ion bound to the membrane-proximal PAS domain of Tlp1-PTPSD. The map was calculated at 1.4-Å resolution and contoured at the 3.0 σ level. The protein side chains that form direct contacts with the ligand are shown. (B) ITC analysis of acetate binding to Tlp1-PTPSD.

C. jejuni Tlp3 (630 Å^3, (Liu et al., 2015)) or the citrate-binding pocket of the PAS domain of CitA (416 Å^3, (Sevvana et al., 2008)). This suggests that if the membrane-proximal PAS domain of Tlp1 recognises any signal or regulatory molecules *in vivo*, they are likely to be of a small size and similar to acetate in their chemical nature.

3.5. The crystal structure suggests indirect mechanism of ligand recognition

Surprisingly, despite the implication of Tlp1 in sensing aspartate, analysis of the high-resolution (1.4 Å) Fourier maps revealed no electron density that could be interpreted as aspartate bound to Tlp1-PTPSD. Indeed, no ligand-binding pocket of a size that is large enough to accommodate an amino acid is present in either of the PAS domains of Tlp1-PTPSD. Furthermore, no binding between Tlp1-PTPSD and aspartate could be detected by isothermal titration calorimetry (ITC, Fig. 5B). This suggests that the likely role of aspartate in the crystallisation was to reduce protein solubility by acidifying the solution. To confirm this, we demonstrated that Tlp1-PTPSD did not crystallise in the absense of aspartate unless the pH of the crystallisation buffer was reduced to 5.2 (equivalent to the acidification effect of 10 mM aspartic acid). Furthermore, this implies that sensing of aspartate by Tlp1-PTPSD is likely to be indirect, *e.g. via* an as yet unidentified periplasmic protein (PBP). Indeed, we note that Tlp1 lacks the consensus motif DXXX(R/K)XWYXXA present in all characterised PTPSDs of MCP receptors that recognise amino acids directly (Liu et al., 2015).

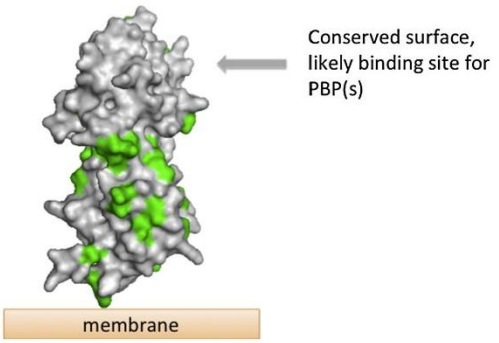


Fig. 7. The locations of strain-to-strain variations (green) on the surface of Tlp1-PTPSD.

PBP-dependent chemoreceptors, albeit with a different sensing domain fold, have been described before. Well-characterised examples of indirect ligand sensing are maltose-binding-protein-mediated recognition of maltose by the *E. coli* aspartic acid receptor Tar (Mowbray and Koshland, 1987) and galactose-binding-protein- and ribose-binding-protein-mediated sugar recognition by the Trg receptor (Aksamit and Koshland, 1974; Hazelbauer and Adler, 1971). Interestingly, our analysis of the distribution of conservative substitutions across 40 different strains with Tlp1 amino-acid sequences showing more than 90% identity to the protein in this study (*C. jejuni* serotype O:2 (strain NCTC 11168)) shows that most strain-to-strain variations map onto the membrane-proximal domain, and that the surface of the membrane-distal domain is highly conserved (Fig. 7). This is consistent with the hypothesis that Tlp1 likely recognises aspartate *via* PBP(s), and that this conserved surface serves as the PBP docking site. The Tlp1-PTPSD structure presented here is anticipated to help in designing experimental approaches to identify the partner PBPs and begin to understand the molecular mechanism of transmembrane signaling mediated by Tlp1 and other homologous chemoreceptors.

Acknowledgements

We thank Victoria Korolik (Griffith University, Australia) for her contribution to the conception of the study and reagents. Part of this research was undertaken on the MX1 and MX2 beamlines of the AS, Victoria, Australia. We thank the AS staff for their assistance with data collection. Mayra A. Machuca is a recipient of a PhD scholarship from the Departamento Admistrativo de Ciencia, Tecnología e Innovación COLCIENCIAS.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jsb.2016.02.019.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. *PHENIX*: a comprehensive python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221.

Aksamit, R.R., Koshland Jr., D.E., 1974. Identification of the ribose binding protein as the receptor for ribose chemotaxis in *Salmonella typhimurium*. Biochemistry 13, 4473–4478.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., Leslie, A.G., 2011. *IMOSFLM*: a new graphical interface for diffraction-image processing with *MOSFLM*. Acta Crystallogr. D Biol. Crystallogr. 67, 271–281.

Boes, J., Nersting, L., Nielsen, E.M., Kranker, S., Enoe, C., Wachmann, H.C., Baggesen, D.L., 2005. Prevalence and diversity of *Campylobacter jejuni* in pig herds on farms with and without cattle or poultry. J. Food Prot. 68, 722–727.

Chen, V.B., Arendall 3rd, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. *MolProbity*: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21.

Dasti, J.I., Tareen, A.M., Lugert, R., Zautner, A.E., Gross, U., 2010. *Campylobacter jejuni*: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. Int. J. Med. Microbiol. 300, 205–211.

Day, C., King, R., Tram, G., Hartley-Tassell, L., Fleetwood, A., Zhulin, I., Korolik, V., 2014. Characterisation of the galactose chemosensory receptor of *Campylobacter jejuni*. In: Proceedings of The Australian Society for Microbiology Annual Scientific Meeting. Australian Society for Microbiology. Abstract #107.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., Liang, J., 2006. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 34, W116–118.

Emsley, P., Cowtan, K., 2004. *Coot*: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Evans, P.R., Murshudov, G.N., 2013. How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69, 1204–1214.

Fernando, U., Biswas, D., Allan, B., Willson, P., Potter, A.A., 2007. Influence of *Campylobacter jejuni fliA, rpoN* and *flgK* genes on colonization of the chicken gut. Int. J. Food Microbiol. 118, 194–200.

Friedman, C., Neimann, J., Wegener, H., Tauxe, R., 2000. Epidemiology of *Campylobacter jejuni* infections in the United States and other industrialized nations. In: Nachamkin, I. (Ed.), Campylobacter. ASM Press, Washington, DC, pp. 121–138.

Glekas, G.D., Foster, R.M., Cates, J.R., Estrella, J.A., Wawrzyniak, M.J., Rao, C.V., Ordal, G.W., 2010. A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. J. Biol. Chem. 285, 1870–1878.

Hartley-Tassell, L.E., Shewell, L.K., Day, C.J., Wilson, J.C., Sandhu, R., Ketley, J.M., Korolik, V., 2010. Identification and characterization of the aspartate chemosensory receptor of *Campylobacter jejuni*. Mol. Microbiol. 75, 710–730.

Hazelbauer, G.L., Adler, J., 1971. Role of the galactose binding protein in chemotaxis of *Escherichia coli* toward galactose. Nature 230, 101–104.

Hendrixson, D.R., Akerley, B.J., DiRita, V.J., 2001. Transposon mutagenesis of *Campylobacter jejuni* identifies a bipartite energy taxis system required for motility. Mol. Microbiol. 40, 214–224.

Henry, J.T., Crosson, S., 2011. Ligand-binding PAS domains in a genomic, cellular, and structural context. Annu. Rev. Microbiol. 65, 261–286.

Hepworth, P.J., Ashelford, K.E., Hinds, J., Gould, K.A., Witney, A.A., Williams, N.J., Leatherbarrow, H., French, N.P., Birtles, R.J., Mendonca, C., Dorrell, N., Wren, B. W., Wigley, P., Hall, N., Winstanley, C., 2011. Genomic variations define divergence of water/wildlife-associated *Campylobacter jejuni* niche specialists from common clonal complexes. Environ. Microbiol. 13, 1549–1560.

Hothorn, M., Dabi, T., Chory, J., 2011. Structural basis for cytokinin recognition by *Arabidopsis thaliana* histidine kinase 4. Nat. Chem. Biol. 7, 766–768.

Hugdahl, M.B., Beery, J.T., Doyle, M.P., 1988. Chemotactic behavior of *Campylobacter jejuni*. Infect. Immun. 56, 1560–1566.

Josenhans, C., Suerbaum, S., 2002. The role of motility as a virulence factor in bacteria. Int. J. Med. Microbiol. 291, 605–614.

King, R.M., Day, C.J., Hartley-Tassell, L.E., Connerton, I.F., Tiralongo, J., McGuckin, M. A., Korolik, V., 2013. Carbohydrate binding and gene expression by *in vitro* and *in vivo* propagated *Campylobacter jejuni* after immunomagnetic separation. J. Basic Microbiol. 53, 240–250.

Korolik, V., Ketley, J.M., 2008. Chemosensory signal transduction pathway of *Campylobacter jejuni*. In: Nachamkin, I. et al. (Eds.), Campylobacter. ASM Press, Washington, DC, pp. 351–366.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. 60, 2256–2268.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. *PROCHECK*: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.

Li, Z., Lou, H., Ojcius, D.M., Sun, A., Sun, D., Zhao, J., Lin, X., Yan, J., 2014. Methyl-accepting chemotaxis proteins 3 and 4 are responsible for *Campylobacter jejuni* chemotaxis and jejuna colonization in mice in response to sodium deoxycholate. J. Med. Microbiol. 63, 343–354.

Liu, Y.C., Machuca, M.A., Beckham, S.A., Gunzburg, M.J., Roujeinikova, A., 2015. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. Acta Crystallogr. D Biol. Crystallogr. 71, 2127–2136.

Machuca, M.A., Liu, Y.C., Roujeinikova, A., 2015. Cloning, expression, refolding, purification and preliminary crystallographic analysis of the sensory domain of the Campylobacter chemoreceptor for aspartate A (CcaA). Acta Crystallogr. F Struct. Biol. Cryst. Commun. 71, 110–113.

Marchant, J., Wren, B., Ketley, J., 2002. Exploiting genome sequence: predictions for mechanisms of *Campylobacter chemotaxis*. Trends Microbiol. 10, 155–159.

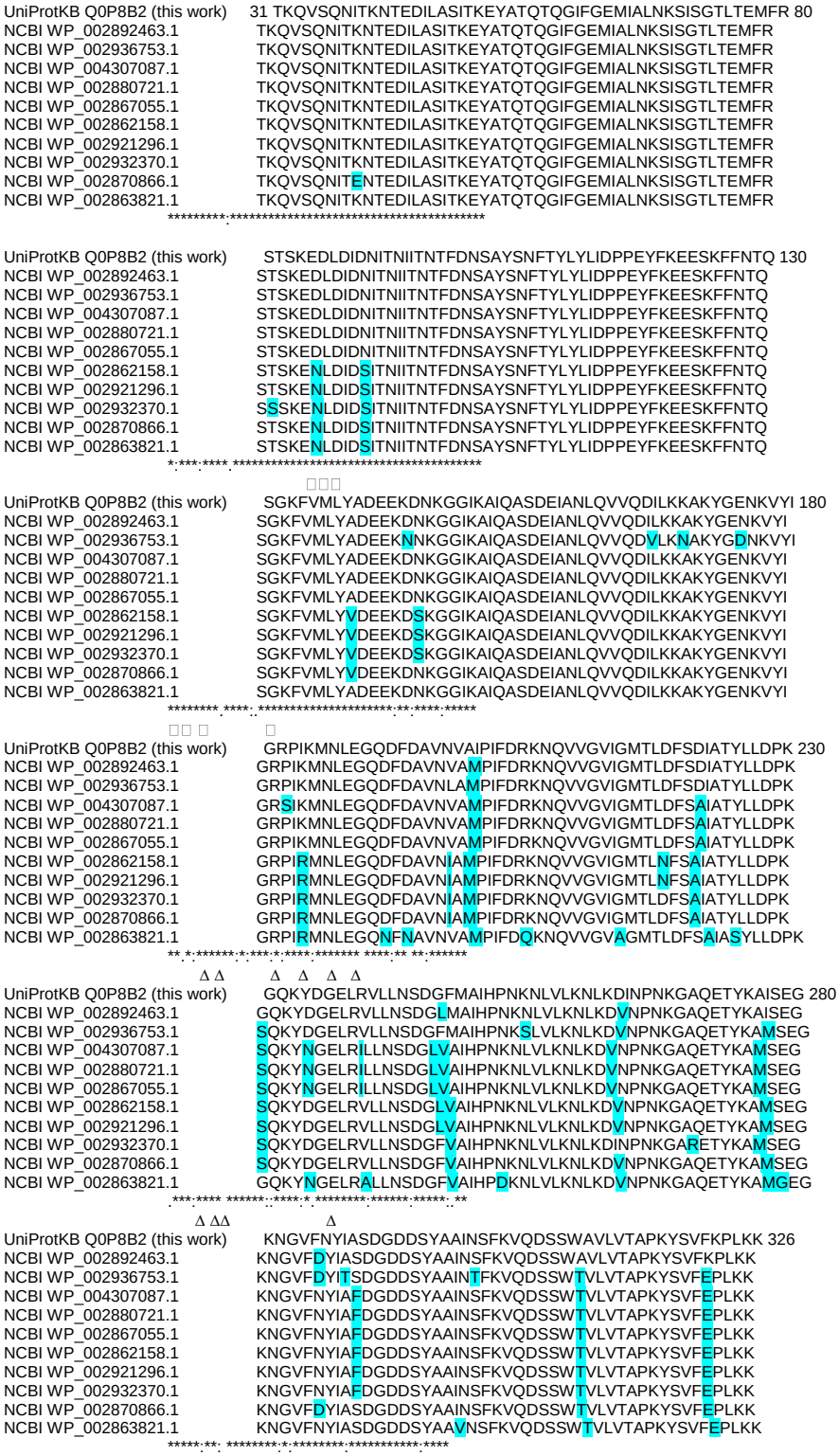
McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., Read, R.J., 2005. Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. 61, 458–464.

Morooka, T., Umeda, A., Amako, K., 1985. Motility as an intestinal colonization factor for *Campylobacter jejuni*. J. Gen Microbiol. 131, 1973–1980.

Mowbray, S.L., Koshland Jr., D.E., 1987. Additive and independent responses in a single receptor: aspartate and maltose stimuli on the tar protein. Cell 50, 171–180.
Neiditch, M.B., Federle, M.J., Miller, S.T., Bassler, B.L., Hughson, F.M., 2005. Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2. Mol. Cell 18, 507–518.
Nishiyama, S., Suzuki, D., Itoh, Y., Suzuki, K., Tajima, H., Hyakutake, A., Homma, M., Butler-Wu, S.M., Camilli, A., Kawagishi, I., 2012. Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. Infect. Immun. 80, 3170–3178.
Oporto, B., Esteban, J.I., Aduriz, G., Juste, R.A., Hurtado, A., 2007. Prevalence and strain diversity of thermophilic campylobacters in cattle, sheep and swine farms. J. Appl. Microbiol. 103, 977–984.
Ortega, A., Amoros, D., Garcia de la Torre, J., 2011. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys. J. 101, 892–898.
Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D., Chillingworth, T., Davies, R.M., Feltwell, T., Holroyd, S., Jagels, K., Karlyshev, A. V., Moule, S., Pallen, M.J., Penn, C.W., Quail, M.A., Rajandream, M.A., Rutherford, K.M., van Vliet, A.H., Whitehead, S., Barrell, B.G., 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature 403, 665–668.
Rahman, H., King, R.M., Shewell, L.K., Semchenko, E.A., Hartley-Tassell, L.E., Wilson, J.C., Day, C.J., Korolik, V., 2014. Characterisation of a multi-ligand binding chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. PLoS Pathog. 10, e1003822.
Rico-Jimenez, M., Munoz-Martinez, F., Garcia-Fontana, C., Fernandez, M., Morel, B., Ortega, A., Ramos, J.L., Krell, T., 2013. Paralogous chemoreceptors mediate chemotaxis towards protein amino acids and the non-protein amino acid gamma-aminobutyrate (GABA). Mol. Microbiol. 88, 1230–1243.
Schmidt-Ott, R., Schmidt, H., Feldmann, S., Brass, F., Krone, B., Gross, U., 2006. Improved serological diagnosis stresses the major role of *Campylobacter jejuni* in triggering Guillain–Barre syndrome. Clin. Vaccine Immunol. 13, 779–783.
Schrödinger, L., 2016. The PyMOL Molecular Graphics System, Version 1.2r3pre.
Sebald, Vernon, 1986. Campylobacteracea. In: Sneath, P.H.A. et al. (Eds.), Bergey's Manual of Systematic Bacteriology. Williams and Wilkins, Baltimore, MD, p. 111.
Sevvana, M., Vijayan, V., Zweckstetter, M., Reinelt, S., Madden, D.R., Herbst-Irmer, R., Sheldrick, G.M., Bott, M., Griesinger, C., Becker, S., 2008. A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA. J. Mol. Biol. 377, 512–523.
Tareen, A.M., Dasti, J.I., Zautner, A.E., Gross, U., Lugert, R., 2010. *Campylobacter jejuni* proteins Cj0952c and Cj0951c affect chemotactic behaviour towards formic acid and are important for invasion of host cells. Microbiology 156, 3123–3135.
Vegge, C.S., Brondsted, L., Li, Y.P., Bang, D.D., Ingmer, H., 2009. Energy taxis drives *Campylobacter jejuni* toward the most favorable conditions for growth. Appl. Environ. Microbiol. 75, 5308–5314.
Wadhams, G.H., Armitage, J.P., 2004. Making sense of it all: bacterial chemotaxis. Nat. Rev. Mol. Cell. Biol. 5, 1024–1037.
Webb, B.A., Hildreth, S., Helm, R.F., Scharf, B.E., 2014. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. Appl. Environ. Microbiol. 80, 3404–3415.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242.
Zautner, A.E., Tareen, A.M., Gross, U., Lugert, R., 2012. Chemotaxis in *Campylobacter jejuni*. Eur. J. Microbiol. Immunol. 2, 24–31.
Zhang, Z., Hendrickson, W.A., 2010. Structural characterization of the predominant family of histidine kinase sensor domains. J. Mol. Biol. 400, 335–353.
Zhou, Y.F., Nan, B., Nan, J., Ma, Q., Panjikar, S., Liang, Y.H., Wang, Y., Su, X.D., 2008. C4-dicarboxylates sensing mechanism revealed by the crystal structures of DctB sensor domain. J. Mol. Biol. 383, 49–61.
Zhulin, I.B., 2001. The superfamily of chemotaxis transducers: from physiology to genomics and back. Adv. Microb. Physiol. 45, 157–198.
Zilbauer, M., Dorrell, N., Wren, B.W., Bajaj-Elliott, M., 2008. *Campylobacter jejuni*-mediated disease pathogenesis: an update. Trans. R. Soc. Trop. Med. Hyg. 102, 123–129.

Supplementary data

Sequence ID



Supplementary Figure S1.

Alignment of 11 representative sequences of Tlp1-PTPSDs from different strains of *C. jejuni*. The subset has been selected from a total of 40 sequences with more than 90% sequence

identity to Tlp1 from *C. jejuni* serotype O:2 (strain NCTC 11168) to show the locations of substitutions (highlighted in colour). The positions of the residues that form the chloride- and acetate-binding sites are shown by open circles and triangles, respectively. http://dx.doi.org/10.1016/j.jsb.2016.02.019

Reference

1 Hugdahl, M. B., Beery, J. T. & Doyle, M. P. Chemotactic behavior of *Campylobacter jejuni*. *Infection and immunity* **56**, 1560-1566 (1988).

2 Sebald & Vernon. Campylobacteracea. *In: Sneath, P.H.A. et al. (Eds.), Bergey's Manual of Systematic Bacteriology. Williams and Wilkins, Baltimore, MD, p. 111.* (1986).

3 Kelly, D. J. in *Campylobacter , Third Edition* (American Society of Microbiology, 2008).

4 Stahl, M., Butcher, J. & Stintzi, A. Nutrient acquisition and metabolism by Campylobacter jejuni. *Front Cell Infect Microbiol* **2**, 5, doi:10.3389/fcimb.2012.00005 (2012).

5 Parkhill, J. *et al.* The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665-668, doi:10.1038/35001088 (2000).

6 Muraoka, W. T. & Zhang, Q. Phenotypic and genotypic evidence for L-fucose utilization by Campylobacter jejuni. *J Bacteriol* **193**, 1065-1075, doi:10.1128/JB.01252-10 (2011).

7 Wright, J. A. *et al.* Metabolite and transcriptome analysis of Campylobacter jejuni in vitro growth reveals a stationary-phase physiological switch. *Microbiology* **155**, 80-94, doi:doi:10.1099/mic.0.021790-0 (2009).

8 Velayudhan, J., Jones, M. A., Barrow, P. A. & Kelly, D. J. L-serine catabolism via an oxygen-labile L-serine dehydratase is essential for colonization of the avian gut by Campylobacter jejuni. *Infect Immun* **72**, 260-268 (2004).

9 Hartley-Tassell, L. E. *et al.* Identification and characterization of the aspartate chemosensory receptor of *Campylobacter jejuni*. *Mol Microbiol* **75**, 710-730, doi:10.1111/j.1365-2958.2009.07010.x (2010).

10 Day, C. *et al.* Characterisation of the galactose chemosensory receptor of *Campylobacter jejuni*. *In: Proceedings of The Australian Society for Microbiology Annual Scientific Meeting. Australian Society for Microbiology.* **Abstract #107.** (2014).

11 Korolik, V. & Ketley, J. in *Campylobacter , Third Edition* (American Society of Microbiology, 2008).

12 Machuca, M. A., Liu, Y. C. & Roujeinikova, A. Cloning, expression, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for aspartate A (CcaA). *Acta crystallographica. Section*

F, Structural biology communications **71**, 110-113, doi:10.1107/S2053230X14027381 (2015).

13 Machuca, M. A., Liu, Y. C., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni* Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition. *J. Struct. Biol.* **194**, 205-213, doi:10.1016/j.jsb.2016.02.019 (2016).

Chapter 4: Structural basis for ligand recognition by dCache ligand binding domain domain of *H. pylori* transducer-like protein C (TlpC)

Preface for Chapter 4

Chemotaxis is essential for *H. pylori*'s ability to swim through the mucus of the stomach towards the epithelium, and survive in the host environment under conditions of constant turnover of the gastric mucosa [1]. In *H. pylori,* chemoeffectors are sensed by four chemoreceptors TlpA, TlpB, TlpC, and TlpD. This pathogen exhibits chemotactic responses to various amino acids [2], cholesterol [3], bicarbonate, bile acids [2], metal ions [4], pH [5], reactive and oxygen species [6]. Although some of these signals have been mapped onto a specific Tlp [4,5,7,8], the mechanism by which most of them are sensed and the chemoreceptor involved is currently unknown.

Among the four chemoreceptors identified in *H. pylori,* TlpC is the least characterised, and its natural ligand remains unknown. The LBD of this receptor shares no homology at the level of amino acid sequence with any of the receptors characterised to date. However, a sequence based search using HHpred [9] predicted secondary structure homology to the LBD of *C. jejuni* Tlp1 and Tlp3 chemoreceptors [10,11] and family 1 histidine kinases [12].

In Chapters Two and Three Tlp1 and Tlp3 receptors were discovered to belong to the dCache family of sensing modules. Therefore, the TlpC predicted homology suggests that its structure may also adopt a fold containing a dCache module with two PAS-Like subdomains. The following set of experiments aimed to structurally characterise the TlpC sensing domain and identify its ligand specificity. Therfore, detailed structural, biophysical and biochemical analysis was performed. The methodology, results and analysis is presented in the journal article format and was published in October 2017 in *Scientifics Reports* [13].

Monash University

Declaration for Thesis Chapter 4

Declaration by candidate

In the case of Chapter 4, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiments, performed experiments, analysed data and wrote manuscript	50

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Kevin S. Johnson	Designed experiments, analysed data and reviewed manuscript	
Yu C. Liu	Designed experiments, analysed data and reviewed manuscript	
David L. Steer	Analysed data	
Karen M. Ottemann	Designed experiments, analysed data, provided intellectual input and reviewed manuscript	
Anna Roujeinikova	Designed experiments, analysed data, wrote manuscript and led the research	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's Signature	Mayra A. Machuca Pérez	**Date** 10/12/17
Main Supervisor's Signature		**Date** 13/12/17

SCIENTIFIC REPORTS

OPEN

Helicobacter pylori chemoreceptor TlpC mediates chemotaxis to lactate

Received: 5 July 2017
Accepted: 6 October 2017
Published online: 26 October 2017

Mayra A. Machuca[1,2], Kevin S. Johnson[3], Yu C. Liu[2], David L. Steer[4], Karen M. Ottemann[3] & Anna Roujeinikova[1,2,5]

It is recently appreciated that many bacterial chemoreceptors have ligand-binding domains (LBD) of the dCACHE family, a structure with two PAS-like subdomains, one membrane-proximal and the other membrane-distal. Previous studies had implicated only the membrane-distal subdomain in ligand recognition. Here, we report the 2.2 Å resolution crystal structure of dCACHE LBD of the *Helicobacter pylori* chemoreceptor TlpC. *H. pylori tlpC* mutants are outcompeted by wild type during stomach colonisation, but no ligands had been mapped to this receptor. The TlpC dCACHE LBD has two PAS-like subdomains, as predicted. The membrane-distal one possesses a long groove instead of a small, well-defined pocket. The membrane-proximal subdomain, in contrast, had a well-delineated pocket with a small molecule that we identified as lactate. We confirmed that amino acid residues making contact with the ligand in the crystal structure—N213, I218 and Y285 and Y249—were required for lactate binding. We determined that lactate is an *H. pylori* chemoattractant that is sensed via TlpC with a $K_D = 155$ µM. Lactate is utilised by *H. pylori*, and our work suggests that this pathogen seeks out lactate using chemotaxis. Furthermore, our work suggests that dCACHE domain proteins can utilise both subdomains for ligand recognition.

Helicobacter pylori is a motile, gram-negative bacterium that infects over 50% of the world's population[1]. *H. pylori* selectively colonises the gastric epithelium and is able to survive in the host stomach for years. Although the majority of the infected people remain asymptomatic, *H. pylori* infection can be associated with a range of gastroduodenal diseases, including gastritis, gastric and duodenal ulcers, and different types of cancer including mucosa-associated lymphoid tissue (MALT) lymphoma and gastric adenocarcinoma[2–4]. Directed motility, or chemotaxis, is important for the ability of *H. pylori* to swim through the highly acidic lumen towards the epithelium and to survive in the host environment under the conditions of constant turnover of the gastric mucosa. Non-motile or non-chemotactic mutants have been shown to be less effective in colonising the gastric mucosa and do not attain full infection compared to the wild type in animal models[5–8].

Chemotaxis allows motile bacteria to sense chemical cues and find optimal environments for growth by, for example, swimming towards favourable chemicals (chemoattractants) and away from harmful ones (repellents). Extracellular chemicals are sensed by chemoreceptors, also termed transducer-like proteins (Tlps). Most of the characterised Tlps are dimeric membrane proteins that comprise an extracytoplasmic ligand-binding domain (LBD), the transmembrane region, the HAMP (**h**istidine kinases, **a**denylyl cyclases, **m**ethyl-accepting protein, and **p**hosphatases) domain and the methyl-accepting (MA) domain (Fig. 1a), the latter transmitting information to a signalling cascade. The signal is relayed through the coupling protein CheW to the histidine protein kinase, CheA, which phosphorylates the response regulator protein, CheY, altering its affinity to the flagellar motors and, as a consequence, the direction (clockwise or counter-clockwise) in which they rotate[9].

The chemotaxis pathway has been extensively studied in *Escherichia coli*[10–12]. Recognition of cue molecules in this bacterium is mediated by five different chemoreceptors[9,12,13]. Four of them contain a periplasmic LBD with

[1]Infection and Immunity Program, Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, 3800, Australia. [2]Department of Microbiology, Monash University, Clayton, Victoria, 3800, Australia. [3]Department of Microbiology and Environmental Toxicology, University of California Santa Cruz, Santa Cruz, CA, 95064, USA. [4]Monash Biomedical Proteomics Facility, Monash University, Clayton, Victoria, 3800, Australia. [5]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, 3800, Australia. Correspondence and requests for materials should be addressed to K.M.O. (email: ottemann@ucsc.edu) or A.R. (email: anna.roujeinikova@monash.edu)

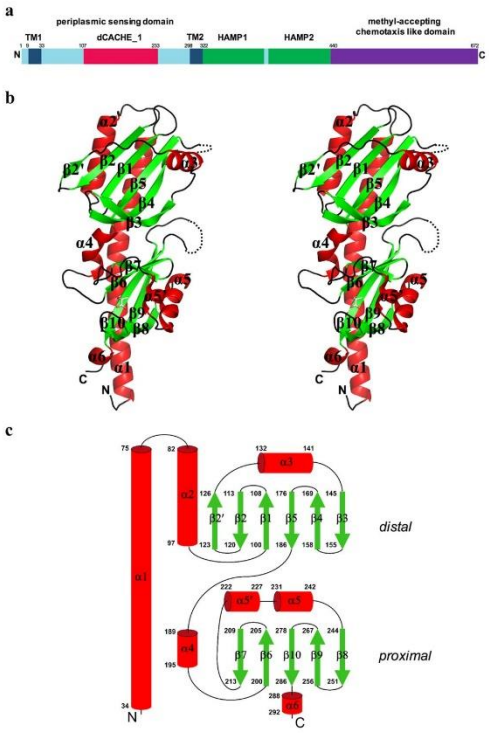


Figure 1. Overall fold of TlpC dCACHE LBD. (**a**) Domain organisation of TlpC, showing LBD location with respect to other structural elements. Transmembrane region (TM, dark blue); dCACHE_1 domain (red), HAMP domains (green); methyl-accepting chemotaxis-like domain (purple). (**b**) Stereo representation of structure of TlpC LBD monomer. (**c**) Topology of secondary structure elements of TlpC LBD. The α-helices are represented by rods and β-strands by arrows. The membrane-distal and membrane-proximal subdomains are labelled.

a 4-helix bundle (4HB) fold. The fifth receptor, Aer, has a cytosolic Per–Arnt–Sim (PAS) LBD and is involved in aerotaxis[14].

Chemoreceptors have been classified according to the size of their LBD into cluster I (~150 amino acids) or cluster II (~250 amino acids)[15]. Much of what is known about bacterial chemoreceptors comes from studies on cluster I chemoreceptors with a 4HB LBD. However, more recent studies have shown that extra-cytoplasmic LBDs of chemoreceptors from different bacteria vary largely in their amino acid sequence and three-dimensional structure[16] and, to date, additional structural families have been identified, including single CACHE (sCACHE)[17–19], helical biomodular (HBM)[20,21] and double CACHE (dCACHE) domains[19,22–25].

In *H. pylori*, four chemoreceptors have been identified based on full genome sequence analysis: TlpA, TlpB, TlpC, and TlpD. TlpD is a soluble, cytoplasmic chemoreceptor that is involved in energy taxis[26] and the repulsion response to reactive oxygen species[27] and acid[28]. TlpA, TlpB, and TlpC are integral membrane proteins[29,30]. TlpA has been linked to recognition of bicarbonate and arginine as attractants[31] and acid as a repellent[28], whilst TlpB has been reported to detect acidic pH[17,30] and the quorum-sensing molecule autoinducer-2 (AI-2)[32] as repellents and direct the chemoattraction response to urea[33]. No signals have been associated with TlpC.

H. pylori additionally exhibits chemotactic responses to various other signals, including amino acids (aspartate, glutamate, asparagine, glutamine, histidine, proline, tyrosine, valine, leucine, serine and glycine)[34], cholesterol[35], bile acids (such as glycocholic, taurocholic, glycodeoxycholic, taurodeoxycholic, glycochenodeoxycholic and taurochenodeoxycholic acid)[34], $ZnCl_2$[36] and $NiCl_2$[31]. However, the recognition of these molecules has not been attributed to a specific chemoreceptor, and the mechanism by which these signals are sensed is currently unknown.

Amongst the four *H. pylori* chemoreceptors, only the periplasmic LBD of TlpB has been structurally characterised[17]. It is a homodimer of sCACHE modules[17,19] – a feature that contrasts to the helical-bundle (4HB) modules of the extensively characterised aspartate receptor Tar from *Salmonella typhimurium*[37], the serine receptor

Tsr from *E. coli*[38], and the McpS chemoreceptor from *Pseudomonas putida*[20]. It is now recognised that the CACHE domain, either in its single sCACHE or double dCACHE form, is the most abundant extracellular sensing domain in prokaryotes, and is commonly found in two-component histidine kinases and chemoreceptors[16,19,22,23,39–41].

TlpC is the least characterised chemoreceptor in *H. pylori*, and its natural ligand was unknown. *H. pylori tlpC* mutants are outcompeted by wild type during stomach colonisation, and TlpC modulates the chemotactic response to acid[29,36]. A BLAST search with the sequence of the sensing domain of TlpC against the structures deposited in the Protein Data Bank (PDB) identified no structural homologues of this domain. However, a pairwise comparison of profile Hidden Markov Models using the HHpred server[42] predicted homology at the level of secondary structure to the sensing domains of family 1 histidine kinases (PDB entries 3lia, 3lib, 3lic, 3lid, 3lif)[43] and chemoreceptors Tlp1 and Tlp3 from *Campylobacter jejuni* (PDB entries 4wy9 and 4xmr)[23,24]. These sensing modules belong to the recently redefined dCACHE_1 structural family[19].

dCACHE domains consist of two structurally similar subdomains that each adopt a PAS-domain-like fold and are arranged in tandem, with one membrane- proximal and the other membrane-distal. dCACHE domain proteins can recognise their signal molecules directly or indirectly. Directly recognised ligands include amino acids[22,23,44–46], pyrimidines[47] and purines[48]. In all previously characterised dCACHE domains, direct sensing involves binding of the signal molecule to the membrane-distal, rather than membrane-proximal, subdomain, and no role for the membrane-proximal subdomain has been determined[22,23,43,44,49].

In this paper, we report the crystal structure of LBD of *H. pylori* TlpC in complex with a small-molecule ligand that co-purified with the protein. The ligand was bound to the membrane-proximal subdomain. Based on the analysis of the electron density maps and the chemical nature of the ligand-binding pocket, we hypothesised and confirmed the ligand to be lactate. The location of the binding site has been validated by mutagenesis. We furthermore verified that lactate acts as an attractant for *H. pylori*, and that TlpC mediates the chemoattractant response. To the best of our knowledge, this is the first example of the dCACHE domain that directly recognises its ligand *via* the membrane-proximal module.

Results

Overall structure of TlpC LBD. The three-dimensional structure of recombinant *H. pylori* TlpC LBD (residues 34–297 plus six additional residues (GIDPFT) at the N-terminus, introduced as an artifact of the cloning procedure), was determined by X-ray crystallography using a single-wavelength anomalous dispersion (SAD) technique to a resolution of 2.2 Å. The TlpC LBD crystals (hereafter referred to as form A) belonged to the space group *C2*, with three molecules in the asymmetric unit related to each other by a three-fold pseudo-symmetry. The coordinates of these molecules were refined independently, and in the final model, they showed very similar backbone conformations that could be superimposed in a pairwise fashion with an overall root mean square deviation (r.m.s.d.) for the C_α atoms of 0.5–0.7 Å. Disordered regions 170–175, 271–274 and 295–297 were not seen in the electron density maps and could not be modelled.

In common with family 1 histidine kinases, the TlpC LBD has a dCACHE fold, and is composed of a membrane-proximal and membrane-distal PAS-like modules folding against the N-terminal and C-terminal halves of a long stalk helix, respectively (Fig. 1b). The TlpC LBD structure comprises six α-helices and 11 β-strands (Fig. 1c). The membrane-distal subdomain (residues 63–186) contains a six-stranded antiparallel β-sheet with the strand order 2′ 2 1 5 4 3, which is flanked on one side by an antiparallel two-helix bundle formed by helix α2 and the C-terminal half of helix α1, and on the other side by helix α3. The membrane-proximal subdomain (residues 34–62, 189–292) contains a five-stranded antiparallel β-sheet with the strand order 7 6 10 9 8. This β-sheet is flanked by an antiparallel two-helix bundle formed by helix α4 and the N-terminal half of helix α1 on one side, and by helices α5 and α5′ on the other side. Finally, an additional helix α6 forms an extension of strand β10 at the C-terminal end of TlpC LBD. The membrane-distal and membrane-proximal subdomains are intimately associated with each other, with a total buried surface area of 1169 Å^2, which is equivalent to 16% of the total buried surface area (BSA) of an individual subdomain.

Analysis of the packing of monomers in the crystal lattice revealed head-to-tail arrangement of molecules, where the membrane-proximal subdomain of one subunit packs against the membrane-distal subdomain of the other, which is not likely to represent a physiologically relevant assembly. To determine the oligomeric state of TlpC LBD in solution, size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) experiments were performed. TlpC LBD eluted as a single monodispersed peak in all conditions tested. The derived molecular weight of 28.8 kDa was close to the value calculated from the amino acid sequence of a monomer (30 kDa). Furthermore, the apparent hydrodynamic radius R_h of the particles in this peak (25 Å) was close the R_h value calculated from the crystal structure of a single TlpC LBD subunit (26 Å). In line with these results, analysis of the crystal packing using the PDBe PISA server (http://www.ebi.ac.uk/pdbe/pisa/) identified no quaternary structure and suggested that TlpC LBD is a monomer in the crystal.

Comparison of TlpC LBD structure to other extracytoplasmic sensing domains. In comparison of TlpC LBD atomic coordinates against structures deposited in the Protein Data Bank using PDBeFold[50], the closest structural similarities were found with the dCACHE_1 sensory modules of chemoreceptors Tlp1 and Tlp3 from *C. jejuni*[23,24], and bacterial family 1 histidine kinases (HK) $HK1_s$-Z8 (*Vibrio parahaemolyticus*), $HK1_s$-Z3 and $HK1_s$-Z2 (*Methanosarcina mazei*)[43]. TlpC LBD structure can be superimposed well over those of Tlp1 (Fig. 2a), $HK1_s$-Z8 (Fig. 2b), $HK1_s$-Z3, $HK1_s$-Z2 and Tlp3 [root-mean-square deviation (r.m.s.d.) of 2.1, 2.2, 2.3, 2.5 and 2.9 Å for 285, 266, 271, 262 and 254 aligned C^α atoms from Tlp1, Z8, Z3, Z2 and Tlp3, respectively], despite the low overall sequence identity (<17%). The dCACHE fold adopted by the TlpC LBD has also been previously observed in sensing domains of chemoreceptor Mlp37[51] and C4-dicarboxylate transport sensory HK DctB from *V. cholerae*[52] (Fig. 2c). Furthermore, this fold is remotely similar to the tandem-PAS fold of LBD of luminescence (lux) system HK LuxQ from *V. harveyi*[53] (Fig. 2c).

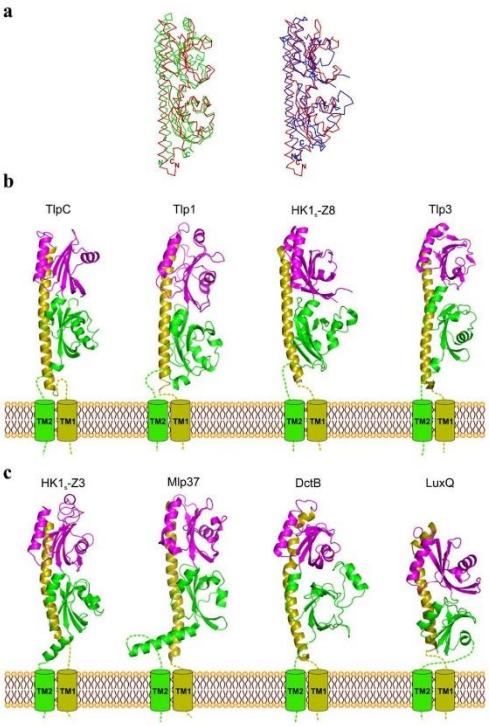


Figure 2. Comparison of dCACHE LBD of TlpC with structures of other dCACHE domains. (**a**,**b**) Structural superposition of TlpC LBD (C_α trace, in red) with dCACHE domains of (**a**) Tlp1 (C_α, in green, PDB entry 4wy9[24]) and (**b**) HK1s-Z8 (C_α, in blue, PDB entry 3lie[43]). (**c**) LBDs of TlpC, Tlp1 (PDB entry 4wy9[24]), HK1s-Z8 (PDB entry 3lie[43]), Tlp3 (PDB entry 4xmq[23]), HK1s-Z3 (PDB entry 3lib[43]), Mlp37 (PDB entry 5ave[51]), DctB (PDB entry 3by9[52]) and LuxQ (PDB entry 1zhh[53]).

Analysis of putative ligand-binding sites. We next examined the putative ligand binding sites, starting with the membrane-distal module. Inspection of the structure of the membrane-distal subdomain around the region implicated in binding of small-molecule ligands in other dCACHE-containing proteins revealed a well-defined groove that runs along the full length of the β-sheet and is flanked on one side by helix α3 and the stretch of amino acid residues connecting β2′ and α3, and on the other side, by the β3-β4-tongue (Fig. 3a and b). This cleft is lined by mostly aliphatic and small hydrophilic residues and has the following approximate dimensions: 30 Å in length, 11 Å in width, and 9 Å in depth. Structural comparisons show that the groove in the membrane-distal domain of TlpC is significantly larger than the small pocket present in Tlp3 and other dCACHE sensing domains that recognise their small-molecule ligands directly[23,24]. Superimposition of the membrane-distal subdomains of TlpC and Tlp3 over 121 C^α atoms (r.m.s.d. of 2.9 Å) shows that helix α3 and the β3-β4-tongue in TlpC are positioned significantly further apart than the equivalent α-helix and β-tongue in Tlp3 (Fig. 3b). The cleft in TlpC appears too large to be a small-molecule-ligand binding site and could hypothetically fit a molecule of the size of a peptide, such as, for example, a loop or a terminal peptide of an-as-yet unidentified PBP.

We then analysed the molecular surface of the membrane-proximal subdomain of TlpC LBD using CASTp[54] with a probe radius of 1.4 Å. We detected a putative ligand-binding pocket with the surface area and solvent-accessible volume of 203 Å and 196 $Å^3$, respectively (Fig. 3a). There was a clear electron density for a non-protein molecule bound in this pocket (Fig. 4). However, its shape did not match any of the components of the purification or crystallisation buffers, which suggested that the ligand trapped in the crystal could be a molecule that was present in the refolding mix or a product of proteolytic degradation of the sample.

Identification of lactate as a ligand for TlpC. To identify the ligand captured by TlpC LBD, the protein was denatured to release the small molecules, and these were analysed by liquid chromatography-electrospray ionisation mass spectrometry (LC-ESI-MS). The negative ionisation mode MS data showed a small peak at $m/z = 89.022$ that was absent in the buffer control (Supplementary Fig. 1). Within the experimental error, this

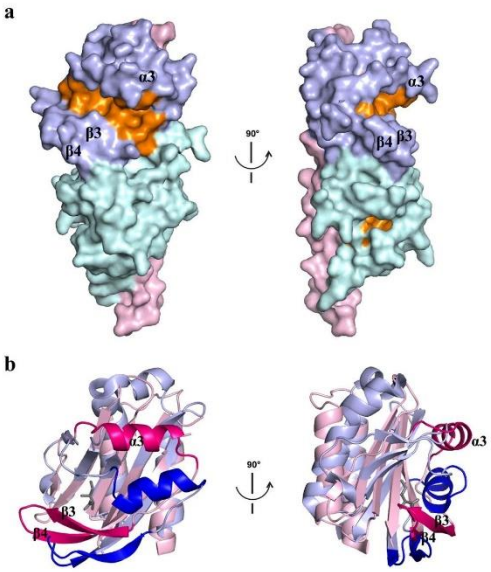


Figure 3. Putative ligand-binding sites of TlpC LBD. (**a**) Molecular surface of TlpC LBD with cavities and pockets coloured orange. The stalk helix is coloured pink, the membrane-distal module – light blue and the membrane-proximal module – cyan. (**b**) Structure superposition of membrane-distal modules of TlpC (pink) and Tlp3 (light blue) highlighting differences in position of helix α3 and β3-β4-tongue (coloured hot pink and blue in TlpC and Tlp3, respectively). Isoleucine bound to the membrane-distal module of Tlp3 is shown as grey sticks.

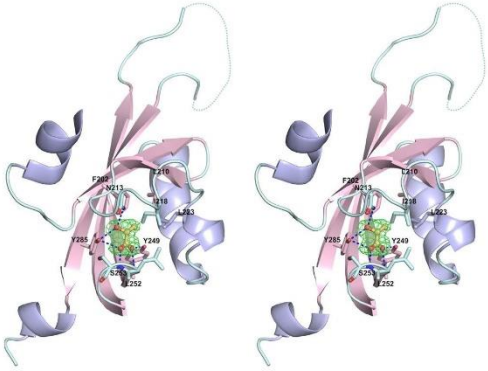


Figure 4. Architecture of ligand-binding site in membrane-proximal module of TlpC dCACHE domain. The (mFo - DFc) σ_A-weighted electron density for lactate is shown in green. The map was calculated at 2.2 Å resolution and contoured at the 3.0 σ level. The lactate molecule is shown in all-atom ball-and-stick representation with C atoms coloured orange. The protein side chains that form direct contacts with lactate are shown in stick representation.

peak matched the chemical formula $C_3H_5O_3$ (m/z 89.024). A search in the PubChem database (https://pubchem.ncbi.nlm.nih.gov/search/) identified 47 different compounds matching this formula. The shape of four of these (lactate, 1,1-dihydroxypropan-2-one, hydron-2-hydroxypropanoate and prop-2-ene-1,1,2-triol) matched the

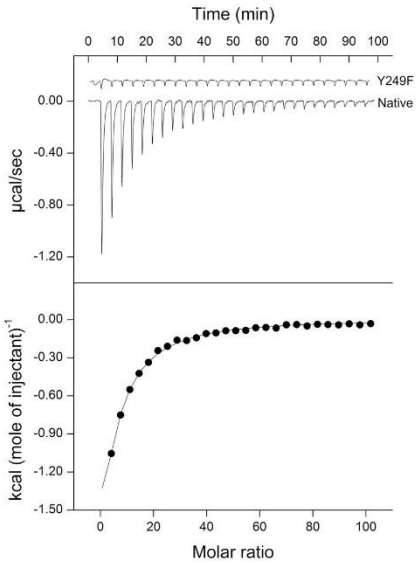


Figure 5. ITC titrations of TlpC LBD and its Y249F variant with lactate. The upper panel shows raw titration data, where each peak corresponds to the injection of 10 µl of 5 mM sodium *L*-lactate into a 1.45-ml reaction cell containing 10 µM protein. The lower panel shows the integrated and dilution-corrected peak areas of the titration plot.

shape of the electron density in the membrane-proximal pocket. As lactate is the only one of these four compounds that is a natural metabolite produced during *E. coli* growth, we hypothesised that during refolding, the lactate was captured by the protein from the cell lysate.

Isothermal titration calorimetry (ITC) measurements confirmed that *L*-lactate binds exothermically to TlpC LBD with an apparent K_D of 155 µM ± 5 µM (Fig. 5). The binding is driven by a favourable enthalpy change ($\Delta H = -20$ kcal mol^{-1}) and is associated with a minor unfavourable entropy change ($T\Delta S = -1.29$ kcal mol^{-1}). This binding appears specific to lactate because no significant heat release or absorption was observed with pyruvate, malate or oxaloacetate, that are chemically similar and metabolically exchangeable with lactate (Supplementary Fig. 2).

Validation of lactate binding site in membrane-proximal module of TlpC dCACHE domain. To establish whether lactate binds to the membrane-distal or membrane-proximal module, we determined the crystal structure of TlpC dCACHE domain co-crystallised with 10 mM *L*-lactate. The co-crystals with lactate were isomorphous to the form A crystals grown with no lactate in the crystallisation mix. Superposition of the protein contents of the two asymmetric units based on the overlap of 767 C_α atoms with an r.m.s.d. of 0.32 Å showed that, within the limit of the experimental error in the coordinates (0.33 Å for the co-crystal with lactate), their structures were essentially identical. Analysis of the electron density maps revealed no lactate binding sites other than the one in the membrane-proximal subdomain. This subdomain contained a lactate molecule bound in a very similar mode to that observed in the form A crystals grown with no added lactate (Fig. 4).

The lactate binding site is located in a pocket formed by residues F202, L210, N213, I218, L223, Y249, L252, S253, and Y285. Calculation of the accessible surface area (ASA) showed that lactate becomes almost completely shielded from the solvent upon binding to TlpC LBD, with 99.5% of its ASA buried by the protein. The carboxyl and hydroxyl groups of lactate form hydrogen bonds with the side chains of N213, Y249 and Y285, and with the main-chain amides of L252 and S253. The TlpC LBD/lactate complex is further stabilised by hydrophobic interactions between the methyl group of lactate and the side chains of F202, L210, I218 and L223 (Fig. 4).

To evaluate the contribution of individual amino acids to the lactate binding, N213, I218 and Y285 were individually replaced with alanine and Y249 with phenylalanine, and the effect of the single-amino acid substitutions was assessed by isotitration calorimetry. Comparison of the circular dichroism spectra of the variants with that of native TlpC LBD showed no significant differences, indicating that the amino acid substitutions did not alter the secondary structure (Fig. 6). ITC measurements demonstrated that each of the N213A, I218A, Y285A and Y249F substitutions abolished the binding of lactate to TlpC LBD (Fig. 4 and Table 1). To further confirm that the membrane-distal subdomain does not bind lactate, the TlpC residues S104, Y151 and K153 – occupying the positions structurally equivalent to the ligand-binding residues Tyr118, Tyr167 and Thr170 in the membrane-distal subdomain of Tlp3 – were individually substituted with alanine. In contrast to the effect on the binding to the

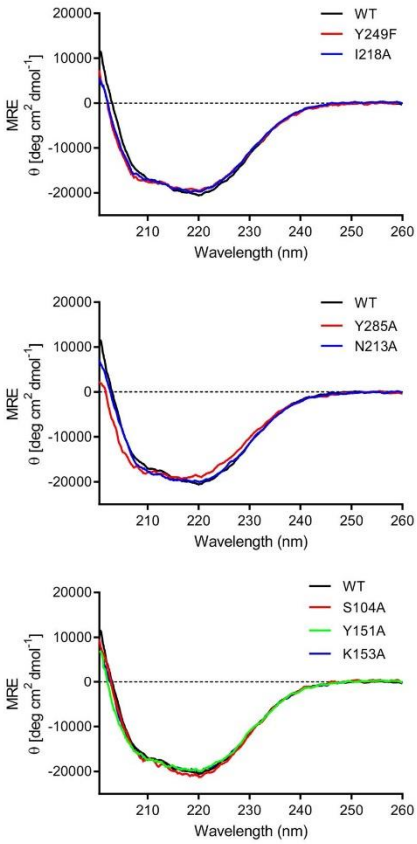


Figure 6. CD spectra of native TlpC LBD (WT) and its S104A, Y151A, K153A, N213A, I218A, Y285A and Y249F variants.

membrane-proximal domain, these substitutions only resulted in 2–3 fold reduction in the affinity to lactate, likely due to partial fold destabilisation rather than loss of interactions with the ligand.

***H. pylori* TlpC mediates positive chemotactic response to lactate.** To test the physiological relevance of the observed specific interaction between TlpC and lactate, we assessed the lactate chemotactic response of *H. pylori*. Although there are isolates of the laboratory *H. pylori* strain 26695 that are motile[55] this strain is prone to motility loss and difficult to use in motility evaluation. We therefore used the human isolate pre-mouse SS1 (PMSS1)[56], which displays a high level of reliable motility, and has been studied for chemotaxis responses in recent publications[28]. The TlpC ligand-binding domain from 26695 and PMSS1 are identical (Supplementary Fig. 3), so we reasoned both proteins would respond similarly to lactate. We assessed whether lactate is an *H. pylori* attractant or repellent using a swimming assay that enumerates flagellar-based bacterial reversals, a common read out for a chemotactic response[57]. In this assay, attractants cause decreased and repellents cause increased direction changes[26,27,32]. Wild-type *H. pylori* showed a significant response to 0.1 mM lactate in this assay, but lost the response at higher concentrations (Fig. 7).

To account for possible chemotactic effects due to pH change upon sodium *L*-lactate treatment of BB10, the pH of the media with and without treatment was assessed. While treatment with 10 mM HCl decreased the pH by more than 1.5 pH units, treatment with any concentration of sodium *L*-lactate only decreased the BB10 pH by less than 0.05 pH units. Furthermore, the pH difference between the highest and lowest amount of sodium *L*-lactate (10 mM and 0.1 mM) was only ~0.01. This analysis suggests that sodium *L*-lactate did not substantially change the medium pH, and thus any effect due to pH change upon sodium *L*-lactate treatment were likely negligible compared to chemotactic effects due to sodium *L*-lactate itself.

Protein	K_D	Enthalpy, ΔH (cal/mol)	Entropy, ΔS (cal/mol/degree)
TlpC LBD native	155.0 ± 5.0 μM	−21,323.3 ± 713.0	−54.1 ± 2.0
TlpC LBD N213A	>3,000	—	—
TlpC LBD I218A	3.1 ± 0.6 mM	−18,145.0 ± 49.0	−49.4 ± 0.2
TlpC LBD Y249A	>3,000	—	—
TlpC LBD Y285F	>3,000	—	—
TlpC LBD F202A	>3,000	—	—
TlpC LBD K223A	>3,000	—	—
TlpC LBD S104A	359.0 ± 3.0 μM	−13,105 ± 318.0	−28.2 ± 1.0
TlpC LBD Y151A	278.5 ± 2.0 μM	−12,040 ± 250.2	−24.1 ± 1.0
TlpC LBD K153A	467.2 ± 5.0 μM	−12,730 ± 345.0	−27.5 ± 1.0

Table 1. Thermodynamic parameters of lactate binding to TlpC LBD and its variants derived from ITC measurements.

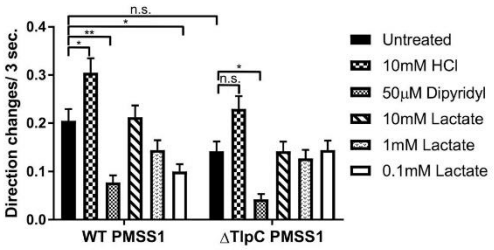


Figure 7. Lactate triggers a TlpC-dependent attractant response in *H. pylori*. Cultures of isogenic wild-type (WT) and Δ*tlpC* (ΔTlpC) *H. pylori* PMSS1 were grown in BB10 overnight, treated with various concentrations of sodium *L*-lactate or control compounds as indicated, and then immediately filmed. Direction changes were enumerated over a 3 second period in at least 200 cells per treatment in two biological replicates. 10 mM HCl or 50 μM dipyridyl serve as known repellent and attractant response controls, respectively. Error bars represent the standard error of the mean. $*p < 0.05$; $**p < 0.01$, comparisons performed using a two-way ANOVA, followed by Tukey's pairwise comparisons ($\alpha = 0.05$). There were no significant differences in the basal behaviour between wild type and its *tlpC* mutant.

We next examined whether TlpC was required for this lactate chemotaxis response. We generated an isogenic null mutant strain lacking *tlpC* (Δ*tlpC*). *tlpC* is part of a single gene operon[58,59], and thus mutations are unlikely to have polar effects. Consistent with this idea of specificity, the Δ*tlpC* mutant retained chemotaxis responses to signals that act through other chemoreceptors: the known repellent HCl that acts through TlpB, TlpA, and TlpD, as well as the attractant dipyridyl that acts through TlpD[27,28,30] (Fig. 7). These responses confirmed that the Δ*tlpC* mutant was not generally chemotaxis defective. Deletion of *tlpC*, however, abolished the chemotactic response to lactate at all tested concentrations (Fig. 7). Additionally, while Δ*tlpC* PMSS1 *H. pylori* displayed fewer reversals on average compared to WT PMSS1, this difference in basal reversal frequency was not significant (Fig. 7). This data thus supports that TlpC is the chemoreceptor for lactate in *H. pylori*.

Discussion

dCACHE domains of bacterial chemoreceptors and histidine kinases consist of two subdomains, membrane-distal and membrane-proximal, each of which could, in principle, harbour a binding site for a small signal or regulatory molecule. In all previously characterised dCACHE domains, including LBDs of Tlp3 from *C. jejuni*[23], Mlp37 from *V. paraehemoliticus*[51], McpB and McpC from *B. subtilis*[22,45], and Mlp24 from *V. cholerae*[60], direct sensing involved binding of the signal molecule to the membrane-distal subdomain. Our analysis of the structural basis of lactate recognition by *H. pylori* chemoreceptor TlpC changes this paradigm regarding the mechanism of sensing by dCACHE domain by providing the first example where direct sensing of the signal molecule is mediated by the ligand binding to the membrane-proximal, rather than membrane-distal, subdomain.

This result has important implications for the conceptual framework of dCACHE-mediated sensing and signal transmission across the membrane. As this and previous studies showed, the membrane-distal and membrane-proximal modules of dCACHE are intimately associated with each other, and are therefore structurally and dynamically coupled[23,24]. For example, we previously demonstrated using X-ray crystallography that, upon binding to an attractant, the dCACHE membrane-distal subdomain of *C. jejuni* Tlp3 closes around the ligand and loses its tight association with the membrane-proximal domain, which, as a result, adopts a more open conformation[23]. The structural coupling of the membrane-distal and membrane-proximal subdomains is

consistent with the finding that signalling across the membrane presumably can be triggered by direct ligand binding to either subdomain – the membrane-proximal subdomain, as in TlpC, or membrane-distal subdomain as in Tlp3, Mlp37, McpB or McpC. Furthermore, one cannot eliminate the possibility that different ligands may signal through the same receptor, with some binding to the membrane-distal and others to the membrane-proximal subdomain. Indeed, all membrane-proximal subdomains of dCACHE sensing domains characterised to date contain a putative small ligand-binding pocket, including those that sense ligands with the membrane-distal domain[23,24,43,51]. For example, the structural study of LBD of *C. jejuni* chemoreceptor Tlp1 revealed an acetate ion bound to the dCACHE membrane-proximal module[24]. Acetate has not yet been found to trigger a chemotaxis response in *C. jejuni*[61]. However, it has been shown to induce either a positive or negative chemotactic response in other species[18,62–64].

Apart from implicating the membrane-proximal, rather than membrane-distal, subdomain in direct sensing of a small molecule ligand, our crystallographic analysis revealed one more difference between the dCACHE domain of TlpC and that of other structurally characterised chemoreceptors of this type – the presence of a long groove in the membrane-distal subdomain, instead of a small well-defined pocket[23,43,51]. This groove might represent a putative binding site for a larger molecule, such as an intermediate ligand-binding protein from the periplasmic binding protein (PBP) family. PBP-mediated sensing is used for chemotaxis in other bacteria, where some PBPs have a function of a primary chemoreceptor that recognises and binds a small molecule in the periplasm, and, in the ligand-bound form, associates with its cognate, membrane-bound transducer-like protein, initiating the signal[65]. There are at least six putative PBPs encoded in the *H. pylori* genome which have been shown to bind autoinducer 2[66] and nickel[67], and proposed to bind other compounds including peptides[68], molybdenum, amino acids, and iron[69]. Whilst the membrane-proximal domain of TlpC mediates direct sensing of lactate, its membrane-distal domain – the shape of which does not imply small-molecule binding – may partner with a PBP to sense some other ligand.

Our analysis of the chemotactic behaviour of wild-type *H. pylori* showed that lactate induced an attractant response in a concentration-dependent manner, and that this response was drastically reduced in a chemotactically competent isogenic Δ*tlpC* mutant, demonstrating that TlpC is the primary chemoreceptor for lactate in *H. pylori*. Within the tested range of concentrations of lactate, the response was strongest at 0.1 mM, detectable at 1.0 mM, and not detectable at 10 mM. Putting the observed TlpC-dependent chemotactic behaviour towards *L*-lactate in the context of the receptor-ligand interactions, we note that the order of the optimal concentration at which lactate is sensed by *H. pylori* as attractant (0.1 mM) is the same as the order of the dissociation constant K_D (0.155 mM) for its binding to the TlpC dCACHE domain. *L*-lactate, secreted by gastric mucous cells, reaches the concentration of 0.3–1 mM in gastric juice[70,71]. Presumably lactate forms a gradient with its highest amount at the source, the cells, but the stomach distribution of lactate is not known. Lactate is known to promote *H. pylori* growth in the stomach[72] and in media that is lacking dextrose, suggesting it can serve as either a carbon or energy source, or both[73]. Metabolically, lactate can be generated by lactate dehydrogenase (LDH) from pyruvic acid as the end product of glycolysis. However, LDH can also catalyse the reverse reaction, converting lactate into pyruvate[74,75]. Thus, if exogenous lactate was imported into *H. pylori*, it could be oxidised into pyruvate, which would then enter the tricarboxylic acid cycle. Alternatively, lactate can donate electrons to NADH and, in turn, to the electron transport chain to enhance proton motive force and bacterial energy levels[76]. In *H. pylori*, the proteins necessary for the import and utilisation of lactate have been identified[69,73], including two lactate permeases and two LDHs[73]. At least one of the lactate utilisation genes has been shown to be required for stomach colonisation, supporting the importance of this process *in vivo*[77].

TlpC mutants have mouse stomach colonisation defects but only when competing with wild type[29]. This phenotype is consistent with the idea that either lactate is limiting and wild type utilises it more efficiently, or that wild type follows a lactate gradient and occupies key niches before the *tlpC* mutant can get there. Thus, the *in vivo* fitness defect observed with isogenic TlpC mutants is consistent with an inability to efficiently access and catabolise lactate *in vivo*[29].

Interestingly, several *H. pylori* lab strains appear to lack TlpC protein expression[27]. One of these strains, G27, has a single base indel that creates a frameshift and results in loss of TlpC expression[27]. Another, B128, also has a *tlpC* frameshift but its gerbil-selected daughter strain, 7.13, has regained TlpC expression[27]. These findings suggest that TlpC is not required for lab growth. The stomach may provide different selective pressures such that TlpC expression is an advantage. In support of this idea, four clinical *H. pylori* isolates analysed all expressed TlpC[27].

The dissociation constant for lactate binding to the TlpC dCACHE domain falls within the middle of the range of values reported for ligand binding by other CACHE domains (*e.g.* 23–356 μM for *Pseudomonas syringae* pv. *actinidiae* PscD LBD[78], 0.6–373 μM for *P. putida* KT2440R McpA LBD[79], 1–1000 μM for *B. subtilis* McpC LBD[22]). Lactate is also a chemoattractant for *Pseudomonas aeruginosa*[18]. *P. aeruginosa* senses lactate *via* an sCACHE domain receptor named McpP. While this receptor has not yet been crystallised, it is known to bind lactate with similar affinity to TlpC, with a K_D of 107 μM[18]. McpP additionally binds acetate, pyruvate, and propionate. McpP is similar to several other sCACHE chemoreceptors, suggesting chemotaxis toward lactate and related C2 and C3 carboxylic acids may be widespread.

Our observation that a higher concentration of lactate (10 mM) did not elicit a positive chemotactic response in *H. pylori* is in agreement with the reports that, at a concentration of 10 mM or above, lactate has an inhibitory effect on *H. pylori* growth[72,80]. The anti-*H. pylori* activity of high levels of exogenous lactate was first observed in co-cultures with lactic acid bacteria (LAB)[81–83]. Lin *et al.*[84] demonstrated that short chain fatty acids (SCFA) (acetic, formic, propionic, butyric and lactic acid) secreted by LAB, and the associated low pH values reduce *H. pylori* viability, with lactic acid exhibiting the strongest inhibitory effect out of all tested SCFA[84,85]. Although the exact mechanism by which high levels of lactic acid exert anti-*H. pylori* activity remains unclear, it is likely a combination of its inhibitory effect on *H. pylori* urease activity and the reduced ability of *H. pylori* to survive at low pH in the absence of urea[80,84–87].

Dataset	Native	K_2PtCl_4	K_2PtCl_6	Co-crystal with 10 mM lactate
Space group	*C*2	*P*321	*P*321	*C*2
a, *b*, *c* (Å)	189.3, 103.2, 61.8	102.7, 102.7, 62.4	102.5, 102.5, 63.0	188.5, 102.6, 61.2
β (°)	98.3			98.5
Resolution range (Å)	30.6–2.2	62.4–3.3	51.4–3.3	30.0–2.5
	(2.31–2.20)	(3.48–3.30)	(3.48–3.30)	(2.60–2.50)
R_{merge}	0.062 (0.280)	0.095 (0.364)	0.096 (0.302)	0.069 (0.341)
Average I/σ(I)	13.3	18.2	18.1	58.5
Completeness (%)	98 (96)	99.9 (99.9)	99.9 (99.9)	99.8 (99.9)
Redundancy	3.6	10.5	10.5	3.6
Anomalous redundancy		5.5	5.4	
Observed reflections	215,101	62,341	62,541	505,592
Unique reflections	59,028	5,941	5,975	40,012

Table 2. X-ray Data collection and processing statistics. Values in parentheses are for the highest resolution shell.

Full-length chemoreceptors function as trimers of dimers[88]. Although the degree and mechanism of the contribution dCACHE LBDs make to oligomerisation *in vivo* remains to be established, previous crystallographic studies on the dCACHE domains of *C. jejuni* Tlp3[23], *M. mazei* $HK1_s$-Z3[43] and *V. cholerae* DctB[52] suggested that they likely dimerise through their stalk helix, with the twofold axis approximately perpendicular to the membrane plane. The dimerisation forces between isolated dCACHE domains are weak, as all domains of this type characterised so far, including that of *H. pylori* TlpC (this study), Tlp1 and Tlp3 from *C. jejuni*[23,24], CtaA and CtaB from *Pseudomonas fluorescens*[89,90], VfcA from *Vibrio fischeri*[91], and PctA from *P. putida*[92], are monomeric in solution. Our analysis showed that the isolated recombinant dCACHE LBD of TlpC is monomeric in the crystal as well. However, the observed structural similarity between LBDs of TlpC, Tlp3, $HK1_s$-Z3 and DctB allows for the possibility that, in the context of the membrane-embedded full-length receptor, TlpC LBD may also dimerise through its stalk helix.

In conclusion, this study reports the first example of the dCACHE type chemoreceptor that directly senses its ligand *via* its membrane-proximal subdomain, and that *H. pylori* seeks out lactate using chemotaxis. It adds to the mounting evidence that dCACHE sensing domains have evolved to recognise their ligands *via* several different direct and indirect mechanisms that may utilise either the membrane-distal, or the membrane-proximal, subdomain, or both. This raises an intriguing question about whether, despite this diversity, different dCACHE sensing domains share a common mechanism of signal transduction across the membrane.

Methods

Site-directed mutagenesis, protein expression, and purification. The expression vectors for single-point variants of TlpC LBD in which S104, Y151, K153, N213, I218 or Y285 were replaced by alanine, and Y249 by phenylalanine, were prepared from a TlpC-expressing plasmid described previously[93]. This plasmid expresses codon-optimised *H. pylori* 26695 TlpC LBD consisting of amino acid residues 34–297 (Fig. 1a). Mutants were created using the QuikChange Mutagenesis Kit (Stratagene). TlpC LBD and its variants were expressed and purified following the previously published procedure[93].

Crystallisation, data collection and structure determination. Form A crystals of TlpC LBD were obtained as described[93]. The crystals grew in space group *C*2 (Table 2) and contained three monomers in the asymmetric unit. Co-crystallisation with 10 mM sodium *L*-lactate under similar conditions produced the crystals of the TlpC LBD/lactate complex that were isomorphous with form A crystals (Table 2). Two platinum derivatives were obtained by soaking the TlpC LBD crystals in either potassium tetrachloroplatinate (1 mM) or potassium hexachloroplatinate (1 mM). The derivative crystals belonged to space group *P*321 (form B) with a monomer in the asymmetric unit (Table 2). Native X-ray diffraction data ($\lambda = 0.95$ Å) and SAD data for the derivatives ($\lambda = 1.07$ Å) were collected on the MX1 and MX2 beamlines of the Australian Synchrotron (AS)[94]. All data were processed with *iMOSFLM*[95] and scaled with AIMLESS[96] from the CCP4 software suite[97] (Table 2).

The two isomorphous SAD data sets were used to locate the platinum sites and calculate the phases for the form B crystals with Autosol[98] from the PHENIX software suite[99]. The resulting phase set (overall figure of merit of 0.30 for data between 51.4 and 3.3 Å) was used to generate an initial partial model for a TlpC LBD monomer with AutoBuild (PHENIX). This model was used for phasing the form A data for the native crystal and the co-crystal with lactate by molecular replacement. Refinement statistics and stereochemistry are given in Table 3. For both models, all the non-glycine residues lie in permitted regions of the Ramachandran plot, with 97% of these in the most favoured regions.

SEC-MALS analysis. The hydrated molecular mass and hydrodynamic radius of TlpC LBD in solution were determined by SEC-MALS. Protein was dialysed against buffer *A* containing 100 mM Tris-HCl pH 8.0 and 150 mM NaCl, and concentrated to 3 mg ml^{-1}. A 100 µl sample was loaded onto a WTC-030S5 SEC column (Wyatt Technology Corporation) pre-equilibrated with the same buffer flowing at 0.4 ml min^{-1}. The eluate was passed through an inline DAWN HELEOS light scattering detector, an Optilab T-rEX differential refractive index

Data set	TlpC native	Co-crystal with 10 mM lactate
Resolution range (Å)	30.6–2.2	30.0–2.5
No. reflections	59,028	40,012
R_{work}/R_{free} [a]	0.182/0.218	0.192/0.251
No. atoms		
Protein	6310	6212
Water	452	66
Lactate	18	18
B-factors		
Average *B* (protein atoms) (Å^2)	46	42
Average *B* (water molecules) (Å^2)	52	37
Average *B* (lactate) (Å^2)	41	41
R.m.s. deviations from ideality		
Bond lengths (Å)	0.017	0.008
Bond angles (°)	1.4	1.1

Table 3. Refinement statistics. [a]The R_{free} was calculated on 5% of the data omitted at random.

Sample	Polydispersity	Molecular weight (kDa)	R_h (nm)
TlpC LBD	1.0	28.8	2.5
TlpC LBD + sodium *L*-lactate	1.0	27	2.6
BSA	1.0	63.9	3.6

Table 4. Dynamic light-scattering results.

detector and a quasi-elastic light scattering detector (WyattQELS, Wyatt Technology Corporation). The experiment was repeated in the presence of 10 mM sodium *L*-lactate. Calculations of the molecular mass and hydrodynamic radius from the intensity of the scattered light and refractive index were performed using ASTRA 6.0 (Wyatt) (Table 4). Theoretical calculations of the hydrodynamic radius from the crystal structure were carried out using HYDROPRO version 10[100].

LC-ESI-MS analysis. Identification of the ligand captured by TlpC LBD was achieved by extracting small molecules from the purified protein and measuring their masses by LC-ESI-MS. TlpC LBD (30 µM in buffer *A*) was unfolded by boiling at 100 °C for 15 min and then pelleted by centrifugation. Buffer *A* subjected to the same procedure was used as a negative control. 200 µl of the supernatant was directly infused into MicrOTOF-Q quadrupole time-of-flight (TOF) mass spectrometer (Bruker Daltonics), and then nebulised and ionised using the Bruker electrospray source. Data were acquired in both positive and negative ion ESI modes over the mass range of 70 to 200 Daltons. The spectra were processed using the Data Analysis software version 3.4 (Bruker Daltonics).

CD analysis. CD spectroscopy was used to compare secondary structure composition of TlpC LBD and its single-point variants. Protein samples (0.1 mg ml^{-1}) were dialysed against buffer *B* (10 mM sodium phosphate pH 7.4, 150 mM NaCl). Far-UV CD spectra were recorded over the wavelength range 200–260 nm at 25 °C with the scan rate of 20 nm min^{-1} using a JASCO J-815 spectropolarimeter. The spectra were measured in triplicate, averaged and smoothed using the Savitzky-Golay algorithm with a radius of 25[101]. Raw data were converted to mean residue ellipticity θ (in deg cm^2 dmol^{-1})[102].

ITC experiments. TlpC LBD and its variants were dialysed against buffer *A*. 5 mM solutions of sodium lactate, sodium pyruvate, sodium malate and sodium oxaloacetate were prepared by dissolving them in the dialysis buffer. Measurements were performed on a MicroCal VP-ITC instrument microcalorimeter (MicroCal) at 25 °C. Protein (10 µM) in a 1.45-ml sample cell was injected with 25 successive 10-µl aliquots of the lactate solution. Binding isotherms were generated by plotting the heat change evolved per injection versus molar ratio of lactate to protein. The data was fitted to a single-site binding model using non-linear least-squares regression (Origin 7, OriginLab, USA), yielding the binding enthalpy ΔH, dissociation constant K_D, and binding entropy ΔH. Each experiment was repeated three times.

Construction of isogenic Δ*tlpC* mutant in *H. pylori* strain PMSS1. The PMSS1 Δ*tlpC* mutant was created by natural transformation of wild-type PMSS1 with 5 µg of Δ*tlpC::cat* SS1 genomic DNA[29,56]. Chloramphenicol-resistant mutants were selected using 10 µg/ml chloramphenicol on Columbia Horse Blood Agar as previously described[29]. Mutation of *tlpC* was confirmed by PCR and western blot.

Chemotaxis assay. Swimming behaviour assays were done with *H. pylori* PMSS1 strains described above grown in Brucella broth (BD BBL/Fisher) with 10% FBS (Life Technologies) (BB10), with shaking, at 37 °C, under microaerobic conditions of 5% O_2, 10% CO_2, balance N_2. Overnight cultures (~OD_{600} 0.25–0.5) were diluted to an OD_{600} of 0.1 in fresh BB10, and then incubated as above until an OD_{600} of 0.12–0.15 was reached. Motile, OD_{600} 0.12–0.15, cultures were treated with sodium *L*-lactate (0.1 mM, 1 mM, 10 mM) or an equal volume of H_2O as an untreated control. As a repellent control, 10 mM HCL was used as done previously[30,32,36]. As an attractant control, 50 μM dipyridyl was used as done previously[27]. Dipyridyl results in fewer directions changes, an attractant response, dependent on chemotaxis in general and TlpD specifically. Dipyridyl induces and attractant response as it counters reactive oxygen species *via* chelation of iron[27]. The pH of BB10 upon treatment was independently assessed using a Denver Instruments pH meter. Cultures were filmed immediately after ligand addition at 400x magnification using a Hamamatsu Digital Camera C4742-95 with the μManager software (Version 1.4.22), mounted on a Nikon Eclipse E600 phase contrast microscope[103] (Supplementary videos 1–12). Videos were relabeled to blind the observer to the strain identity. For each sample, >100 3-s-long bacterial tracks from two independent cultures were analysed manually to identify stops followed by direction changes and to calculate the average number of direction changes in 3 s. Statistical analysis of the data for treated versus untreated samples was performed using a Student's t-test.

PDB submission codes. The atomic coordinates and structure factors of the TlpC LBD/lactate complex obtained at 2.2 Å resolution have been deposited in the Protein Data Bank (http://www.rcsb.org) under accession code 5wbf.

References

1. Hunt, R. H. *et al. Helicobacter pylori* in developing countries. World Gastroenterology Organisation Global Guideline. *J. Gastrointestin. Liver Dis.* **20**, 299–304 (2011).
2. Uemura, N. *et al. Helicobacter pylori* infection and the development of gastric cancer. *N. Engl. J. Med.* **345**, 784–789 (2001).
3. Peek, R. M. Jr. & Blaser, M. J. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nat. Rev. Cancer* **2**, 28–37 (2002).
4. Marshall, B. J. & Warren, J. R. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* **1**, 1311–1315 (1984).
5. Ottemann, K. M. & Lowenthal, A. C. *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. *Infect. Immun.* **70**, 1984–1990 (2002).
6. Terry, K., Williams, S. M., Connolly, L. & Ottemann, K. M. Chemotaxis plays multiple roles during *Helicobacter pylori* animal infection. *Infect. Immun.* **73**, 803–811 (2005).
7. Keilberg, D. & Ottemann, K. M. How *Helicobacter pylori* senses, targets and interacts with the gastric epithelium. *Environ. Microbiol.* **18**, 791–806 (2016).
8. Foynes, S. *et al. Helicobacter pylori* possesses two CheY response regulators and a histidine kinase sensor, CheA, which are essential for chemotaxis and colonization of the gastric mucosa. *Infect. Immun.* **68**, 2016–2023 (2000).
9. Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem. Sci.* **33**, 9–19 (2008).
10. Wuichet, K., Alexander, R. P. & Zhulin, I. B. Comparative genomic and protein sequence analyses of a complex system controlling bacterial chemotaxis. *Methods Enzymol.* **422**, 1–31 (2007).
11. Miller, L. D., Russell, M. H. & Alexandre, G. Diversity in bacterial chemotactic responses and niche adaptation. *Adv. Appl. Microbiol.* **66**, 53–75 (2009).
12. Wadhams, G. H. & Armitage, J. P. Making sense of it all: bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.* **5**, 1024–1037 (2004).
13. Yeh, J. I. *et al.* High-resolution structures of the ligand binding domain of the wild-type bacterial aspartate receptor. *J. Mol. Biol.* **262**, 186–201 (1996).
14. Bibikov, S. I., Biran, R., Rudd, K. E. & Parkinson, J. S. A signal transducer for aerotaxis in. *Escherichia coli. J. Bacteriol.* **179**, 4075–4079 (1997).
15. Lacal, J., Garcia-Fontana, C., Munoz-Martinez, F., Ramos, J. L. & Krell, T. Sensing of environmental signals: classification of chemoreceptors according to the size of their ligand binding regions. *Environ. Microbiol.* **12**, 2873–2884 (2010).
16. Bardy, S. L., Briegel, A., Rainville, S. & Krell, T. Recent advances and future prospects in bacterial and archaeal locomotion and signal transduction. *J. Bacteriol*, https://doi.org/10.1128/JB.00203-17 (2017).
17. Goers Sweeney, E. *et al.* Structure and proposed mechanism for the pH-sensing *Helicobacter pylori* chemoreceptor TlpB. *Structure* **20**, 1177–1188 (2012).
18. Garcia, V. *et al.* Identification of a Chemoreceptor for C2 and C3 Carboxylic Acids. *Appl. Environ. Microbiol.* **81**, 5449–5457 (2015).
19. Upadhyay, A. A., Fleetwood, A. D., Adebali, O., Finn, R. D. & Zhulin, I. B. Cache Domains That are Homologous to, but Different from PAS Domains Comprise the Largest Superfamily of Extracellular Sensors in Prokaryotes. *PLoS Comput. Biol.* **12**, e1004862; 1004810.1001371/journal.pcbi.1004862, (2016).
20. Pineda-Molina, E. *et al.* Evidence for chemoreceptors with bimodular ligand-binding regions harboring two signal-binding sites. *Proc. Natl. Acad. Sci. USA* **109**, 18926–18931 (2012).
21. Martin-Mora, D. *et al.* McpQ is a specific citrate chemoreceptor that responds preferentially to citrate/metal ion complexes. *Environ. Microbiol.* **18**, 3284–3295 (2016).
22. Glekas, G. D. *et al.* The *Bacillus subtilis* chemoreceptor McpC senses multiple ligands using two discrete mechanisms. *J. Biol. Chem.* **287**, 39412–39418 (2012).
23. Liu, Y. C., Machuca, M. A., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 2127–2136 (2015).
24. Machuca, M. A., Liu, Y. C., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni* Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition. *J. Struct. Biol.* **194**, 205–213 (2016).
25. Matilla, M. A. & Krell, T. Chemoreceptor-based signal sensing. *Curr. Opin. Biotechnol.* **45**, 8–14 (2017).
26. Schweinitzer, T. *et al.* Functional characterization and mutagenesis of the proposed behavioral sensor TlpD of *Helicobacter pylori*. *J. Bacteriol.* **190**, 3244–3255 (2008).
27. Collins, K. D. *et al.* The *Helicobacter pylori* CZB Cytoplasmic Chemoreceptor TlpD Forms an Autonomous Polar Chemotaxis Signaling Complex That Mediates a Tactic Response to Oxidative Stress. *J. Bacteriol.* **198**, 1563–1575 (2016).
28. Huang, J. Y., Goers Sweeney, E., Guillemin, K. & Amieva, M. R. Multiple Acid Sensors Control *Helicobacter pylori* Colonization of the Stomach. *PLoS Pathog.* **13**, e1006118, 10.1001371/journal.ppat.1006118 (2017).

29. Andermann, T. M., Chen, Y. T. & Ottemann, K. M. Two predicted chemoreceptors of *Helicobacter pylori* promote stomach infection. *Infect. Immun.* **70**, 5877–5881 (2002).
30. Croxen, M. A., Sisson, G., Melano, R. & Hoffman, P. S. The *Helicobacter pylori* chemotaxis receptor TlpB (HP0103) is required for pH taxis and for colonization of the gastric mucosa. *J. Bacteriol.* **188**, 2656–2665 (2006).
31. Cerda, O., Rivas, A. & Toledo, H. *Helicobacter pylori* strain ATCC700392 encodes a methyl-accepting chemotaxis receptor protein (MCP) for arginine and sodium bicarbonate. *FEMS Microbiol. Lett.* **224**, 175–181 (2003).
32. Rader, B. A. *et al. Helicobacter pylori* perceives the quorum-sensing molecule AI-2 as a chemorepellent via the chemoreceptor TlpB. *Microbiology* **157**, 2445–2455 (2011).
33. Huang, J. Y. *et al.* Chemodetection and Destruction of Host Urea Allows *Helicobacter pylori* to Locate the Epithelium. *Cell Host Microbe* **18**, 147–156 (2015).
34. Worku, M. L., Karim, Q. N., Spencer, J. & Sidebotham, R. L. Chemotactic response of *Helicobacter pylori* to human plasma and bile. *J. Med. Microbiol.* **53**, 807–811 (2004).
35. Wunder, C. *et al.* Cholesterol glucosylation promotes immune evasion by *Helicobacter pylori*. *Nat. Med.* **12**, 1030–1038 (2006).
36. Sanders, L., Andermann, T. M. & Ottemann, K. M. A supplemented soft agar chemotaxis assay demonstrates the *Helicobacter pylori* chemotactic response to zinc and nickel. *Microbiology* **159**, 46–57 (2013).
37. Milburn, M. V. *et al.* Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* **254**, 1342–1347 (1991).
38. Kim, K. K., Yokota, H. & Kim, S. H. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* **400**, 787–792 (1999).
39. Reinelt, S., Hofmann, E., Gerharz, T., Bott, M. & Madden, D. R. The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular PAS domain. *J. Biol. Chem.* **278**, 39189–39196 (2003).
40. Pappalardo, L. *et al.* The NMR structure of the sensory domain of the membranous two-component fumarate sensor (histidine protein kinase) DcuS of *Escherichia coli*. *J. Biol. Chem.* **278**, 39185–39188 (2003).
41. Neiditch, M. B. *et al.* Ligand-induced asymmetry in histidine sensor kinase complex regulates quorum sensing. *Cell* **126**, 1095–1108 (2006).
42. Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–248 (2005).
43. Zhang, Z. & Hendrickson, W. A. Structural characterization of the predominant family of histidine kinase sensor domains. *J. Mol. Biol.* **400**, 335–353 (2010).
44. McKellar, J. L., Minnell, J. J. & Gerth, M. L. A high-throughput screen for ligand binding reveals the specificities of three amino acid chemoreceptors from *Pseudomonas syringae* pv. *actinidiae*. *Mol. Microbiol.* **96**, 694–707 (2015).
45. Glekas, G. D. *et al.* A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. *J. Biol. Chem.* **285**, 1870–1878 (2010).
46. Webb, B. A., Hildreth, S., Helm, R. F. & Scharf, B. E. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Appl. Environ. Microbiol.* **80**, 3404–3415 (2014).
47. Liu, X., Wood, P. L., Parales, J. V. & Parales, R. E. Chemotaxis to pyrimidines and identification of a cytosine chemoreceptor in *Pseudomonas putida*. *J. Bacteriol.* **191**, 2909–2916 (2009).
48. Fernandez, M., Morel, B., Corral-Lugo, A. & Krell, T. Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol. Microbiol.* **99**, 34–42 (2016).
49. Zhou, Y. F. *et al.* C4-dicarboxylates sensing mechanism revealed by the crystal structures of DctB sensor domain. *J. Mol. Biol.* **383**, 49–61 (2008).
50. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268 (2004).
51. Nishiyama, S. *et al.* Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci. Rep.* **6**, 20866, https://doi.org/10.21038/srep20866 (2016).
52. Cheung, J. & Hendrickson, W. A. Crystal structures of C4-dicarboxylate ligand complexes with sensor domains of histidine kinases DcuS and DctB. *J. Biol. Chem.* **283**, 30256–30265 (2008).
53. Neiditch, M. B., Federle, M. J., Miller, S. T., Bassler, B. L. & Hughson, F. M. Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2. *Mol. Cell* **18**, 507–518 (2005).
54. Dundas, J. *et al.* CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, W116–118 (2006).
55. Josenhans, C., Eaton, K. A., Thevenot, T. & Suerbaum, S. Switching of flagellar motility in *Helicobacter pylori* by reversible length variation of a short homopolymeric sequence repeat in fliP, a gene encoding a basal body protein. *Infect. Immun.* **68**, 4598–4603 (2000).
56. Arnold, I. C. *et al.* Tolerance rather than immunity protects from *Helicobacter pylori*-induced gastric preneoplasia. *Gastroenterology* **140**, 199–209 (2011).
57. Terry, K., Go, A. C. & Ottemann, K. M. Proteomic mapping of a suppressor of non-chemotactic cheW mutants reveals that *Helicobacter pylori* contains a new chemotaxis protein. *Mol. Microbiol.* **61**, 871–882 (2006).
58. Sharma, C. M. *et al.* The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**, 250–255 (2010).
59. Lertsethtakarn, P., Draper, J. & Ottemann, K. M. Chemotactic Signal Transduction in *Helicobacter pylori* in *Two-Component Systems in Bacteria* (eds R. Gross & D. Beier) Ch. 17, 355–370 (Caister Academic Press, 2012).
60. Nishiyama, S. *et al.* Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infect. Immun.* **80**, 3170–3178 (2012).
61. Hugdahl, M. B., Beery, J. T. & Doyle, M. P. Chemotactic behavior of *Campylobacter jejuni*. *Infect. Immun.* **56**, 1560–1566 (1988).
62. Repaske, D. R. & Adler, J. Change in intracellular pH of *Escherichia coli* mediates the chemotactic response to certain attractants and repellents. *J. Bacteriol.* **145**, 1196–1208 (1981).
63. Kihara, M. & Macnab, R. M. Cytoplasmic pH mediates pH taxis and weak-acid repellent taxis of bacteria. *J. Bacteriol.* **145**, 1209–1221 (1981).
64. Hamblin, P. A., Maguire, B. A., Grishanin, R. N. & Armitage, J. P. Evidence for two chemosensory pathways in *Rhodobacter sphaeroides*. *Mol. Microbiol.* **26**, 1083–1096 (1997).
65. Bi, S. & Lai, L. Bacterial chemoreceptors and chemoeffectors. *Cell Mol. Life Sci.* **72**, 691–708 (2015).
66. Anderson, J. K. *et al.* Chemorepulsion from the Quorum Signal Autoinducer-2 Promotes *Helicobacter pylori* Biofilm Dispersal. *MBio* **6**, e00379; https://doi.org/10.01128/mBio.00379-00315 (2015).
67. Shaik, M. M., Cendron, L., Salamina, M., Ruzzene, M. & Zanotti, G. *Helicobacter pylori* periplasmic receptor CeuE (HP1561) modulates its nickel affinity via organic metallophores. *Mol. Microbiol.* **91**, 724–735 (2014).
68. Weinberg, M. V. & Maier, R. J. Peptide transport in *Helicobacter pylori*: roles of dpp and opp systems and evidence for additional peptide transporters. *J. Bacteriol.* **189**, 3392–3402 (2007).
69. Tomb, J. F. *et al.* The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539–547 (1997).
70. Piper, D. W., Fenton, B. H. & Goodman, L. R. Lactic, pyruvic, citric, and uric acid and urea content of human gastric juice. *Gastroenterology* **53**, 42–48 (1967).

71. Armstrong, C. P., Dent, D. M., Berman, P. & Aitken, R. J. The relationship between gastric carcinoma and gastric juice lactate (L + D) and lactate dehydrogenase. *Am. J. Gastroenterol.* **79**, 675–678 (1984).
72. Takahashi, T. *et al.* L-lactic acid secreted from gastric mucosal cells enhances growth of *Helicobacter pylori*. *Helicobacter* **12**, 532–540 (2007).
73. Iwatani, S. *et al.* Identification of the genes that contribute to lactate utilization in *Helicobacter pylori*. *PLoS One* **9**, e103506, 10.101371/journal.pone.0103506 (2014).
74. Garvie, E. I. Bacterial lactate dehydrogenases. *Microbiol. Rev.* **44**, 106–139 (1980).
75. Everse, J. & Kaplan, N. O. Lactate dehydrogenases: structure and function. *Adv. Enzymol. Relat. Areas Mol. Biol.* **37**, 61–133 (1973).
76. Smith, H., Yates, E. A., Cole, J. A. & Parsons, N. J. Lactate stimulation of gonococcal metabolism in media containing glucose: mechanism, impact on pathogenicity, and wider implications for other pathogens. *Infect. Immun.* **69**, 6565–6572 (2001).
77. Baldwin, D. N. *et al.* Identification of *Helicobacter pylori* genes that contribute to stomach colonization. *Infect. Immun.* **75**, 1005–1016 (2007).
78. Brewster, J. L. *et al.* Structural basis for ligand recognition by a Cache chemosensory domain that mediates carboxylate sensing in *Pseudomonas syringae*. *Sci. Rep.* **6**, 35198 (2016).
79. Corral-Lugo, A. *et al.* Assessment of the contribution of chemoreceptor-based signaling to biofilm formation. *Environ. Microbiol.* https://doi.org/10.1111/1462-2920.13170 (2015).
80. Aiba, Y., Suzuki, N., Kabir, A. M., Takagi, A. & Koga, Y. Lactic acid-mediated suppression of *Helicobacter pylori* by the oral administration of *Lactobacillus salivarius* as a probiotic in a gnotobiotic murine model. *Am. J. Gastroenterol.* **93**, 2097–2101 (1998).
81. Bhatia, S. J., Kochar, N., Abraham, P., Nair, N. G. & Mehta, A. P. *Lactobacillus acidophilus* inhibits growth of *Campylobacter pylori in vitro*. *J. Clin. Microbiol.* **27**, 2328–2330 (1989).
82. Canducci, F. *et al.* Probiotics and *Helicobacter pylori* eradication. *Dig. Liver Dis.* **34**(Suppl 2), S81–83 (2002).
83. Gotteland, M., Brunser, O. & Cruchet, S. Systematic review: are probiotics useful in controlling gastric colonization by *Helicobacter pylori*? *Aliment. Pharmacol. Ther.* **23**, 1077–1086 (2006).
84. Lin, W. H. *et al.* Anti-*Helicobacter pylori* activity of fermented milk with lactic acid bacteria. *J. Sci. Food Agric.* **91**, 1424–1431 (2011).
85. Midolo, P. D., Lambert, J. R., Hull, R., Luo, F. & Grayson, M. L. *In vitro* inhibition of *Helicobacter pylori* NCTC 11637 by organic acids and lactic acid bacteria. *J. Appl. Bacteriol.* **79**, 475–479 (1995).
86. Sgouras, D. *et al.* *In vitro* and *in vivo* inhibition of *Helicobacter pylori* by *Lactobacillus casei* strain Shirota. *Appl. Environ. Microbiol.* **70**, 518–526 (2004).
87. Lin, W. H. *et al.* Antagonistic activity of spent culture supernatants of lactic acid bacteria against *Helicobacter pylori* growth and infection in human gastric epithelial AGS cells. *J. Food Sci.* **74**, M225–230 (2009).
88. Briegel, A. *et al.* New insights into bacterial chemoreceptor array structure and assembly from electron cryotomography. *Biochemistry* **53**, 1575–1585 (2014).
89. Ud-Din, A. I. & Roujeinikova, A. Cloning, purification, crystallization and X-ray crystallographic analysis of the periplasmic sensing domain of *Pseudomonas fluorescens* chemotactic transducer of amino acids type A (CtaA). *Biosci. Trends* **10**, 320–324 (2016).
90. Salah Ud-Din, A. I. M. & Roujeinikova, A. The periplasmic sensing domain of *Pseudomonas fluorescens* chemotactic transducer of amino acids type B (CtaB): Cloning, refolding, purification, crystallization, and X-ray crystallographic analysis. *Biosci. Trends* **11**, 229–234 (2017).
91. Salah Ud-Din, A. I. & Roujeinikova, A. The periplasmic sensing domain of *Vibrio fischeri* chemoreceptor protein A (VfcA): cloning, purification and crystallographic analysis. *Acta Crystallogr. F Struct. Biol. Commun.* **72**, 382–385 (2016).
92. Rico-Jimenez, M. *et al.* Paralogous chemoreceptors mediate chemotaxis towards protein amino acids and the non-protein amino acid gamma-aminobutyrate (GABA). *Mol. Microbiol.* **88**, 1230–1243 (2013).
93. Liu, Y. C. & Roujeinikova, A. Expression, refolding, purification and crystallization of the sensory domain of the TlpC chemoreceptor from *Helicobacter pylori* for structural studies. *Protein Expr. Purif.* **107**, 29–34 (2015).
94. McPhillips, T. M. *et al.* Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* **9**, 401–406 (2002).
95. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).
96. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
97. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
98. Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 582–601 (2009).
99. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
100. Ortega, A., Amoros, D. & Garcia de la Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898 (2011).
101. Savitzky, A. & Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **36**, 1627–1639 (1964).
102. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2006).
103. Edelstein, A. D. *et al.* Advanced methods of microscope control using μManager software. *J. Biol. Methods* **1**, e10, https://doi.org/10.14440/jbm.12014.14436 (2014).

Acknowledgements

We thank Dr. Danuta Maksel and Dr. Robyn Gray at the Monash Macromolecular Crystallisation Facility for their assistance in setting up automated crystallization screens. We additionally thank Dr. Susan Williams (UC Santa Cruz) for early work on lactate chemotaxis. Part of this research was undertaken on the MX1 and MX2 beamlines of the AS, Victoria, Australia. We thank the AS staff for their assistance with data collection. The described project was supported by a National Institutes of Health National Institute of Allergy and Infectious Disease (NIAID) grant number RO1AI116946 (to K.M.O.). Mayra A. Machuca is indebted to the Departamento Admistrativo de Ciencia, Tecnología e Innovación COLCIENCIAS for a doctoral scholarship. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author Contributions

A.R., Y.C.L. and K.M.O. conceived the study. M.A.M., K.S.J., Y.C.L., D.L.S. performed the experiments. All authors analyzed the data. M.A.M., K.S.J., Y.C.L. and A.R. prepared the Figures. M.A.M., Y.C.L. and A.R. drafted the manuscript. K.S.J. and K.M.O. edited the manuscript. All authors reviewed and approved the final version.

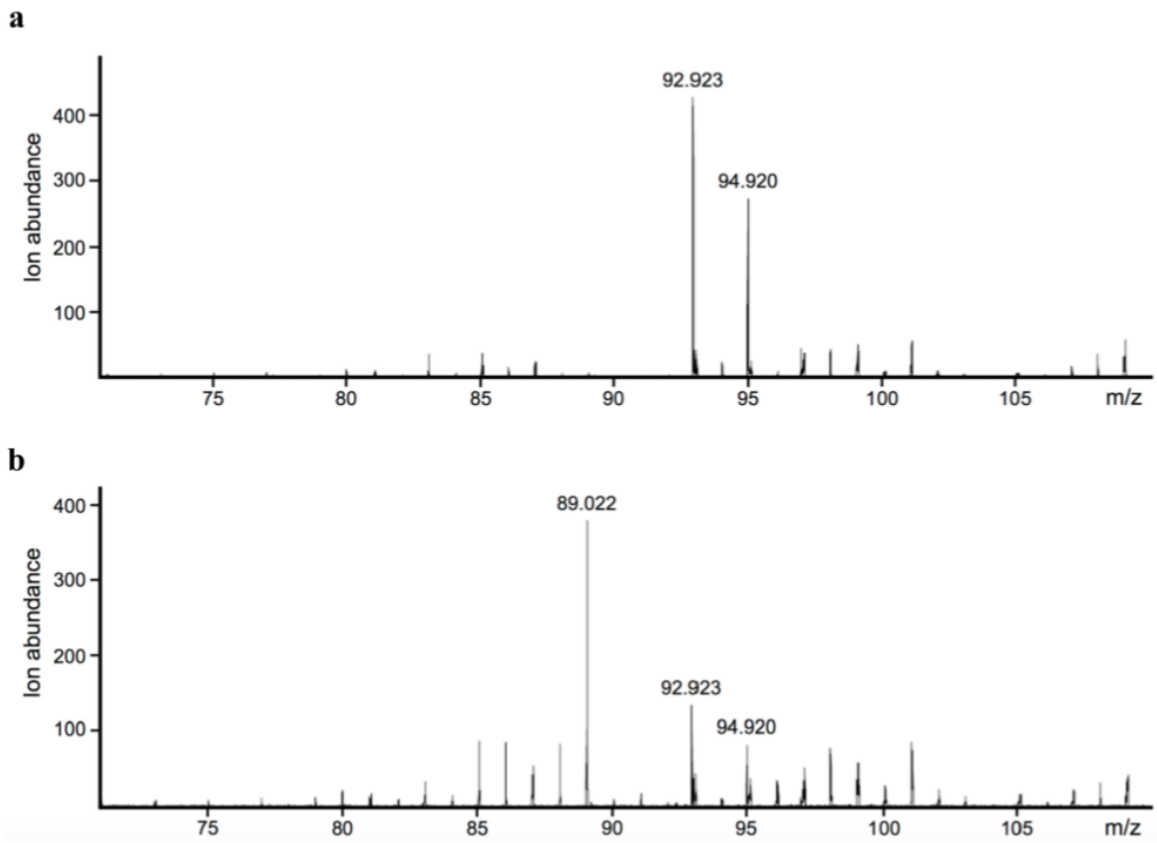
Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-14372-2.

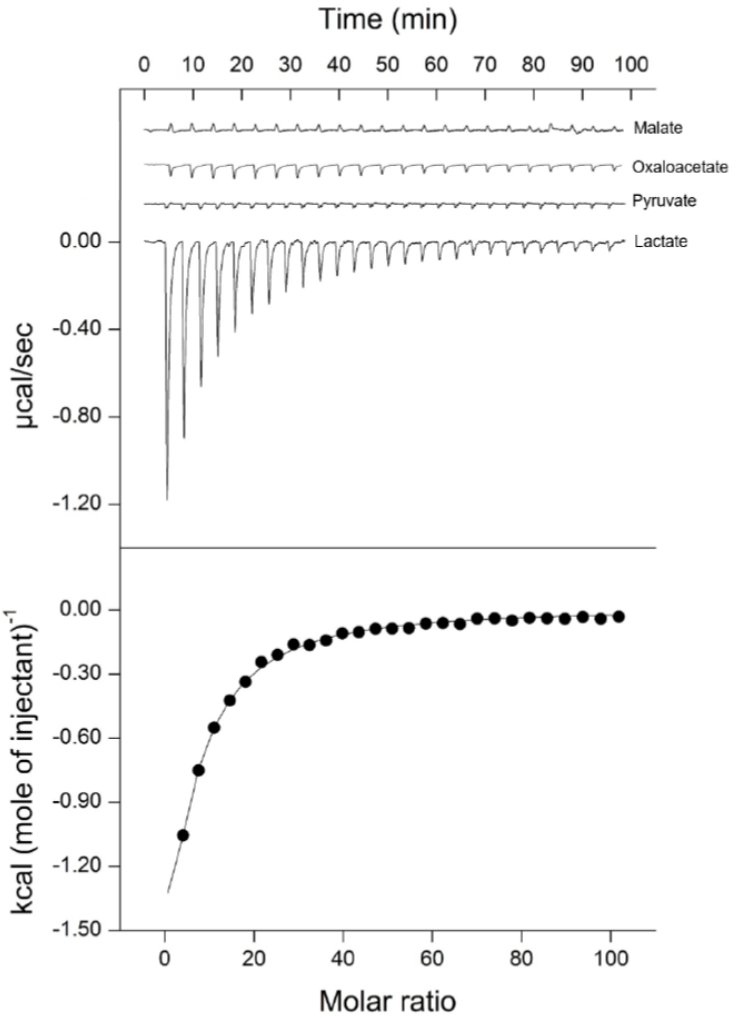
Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

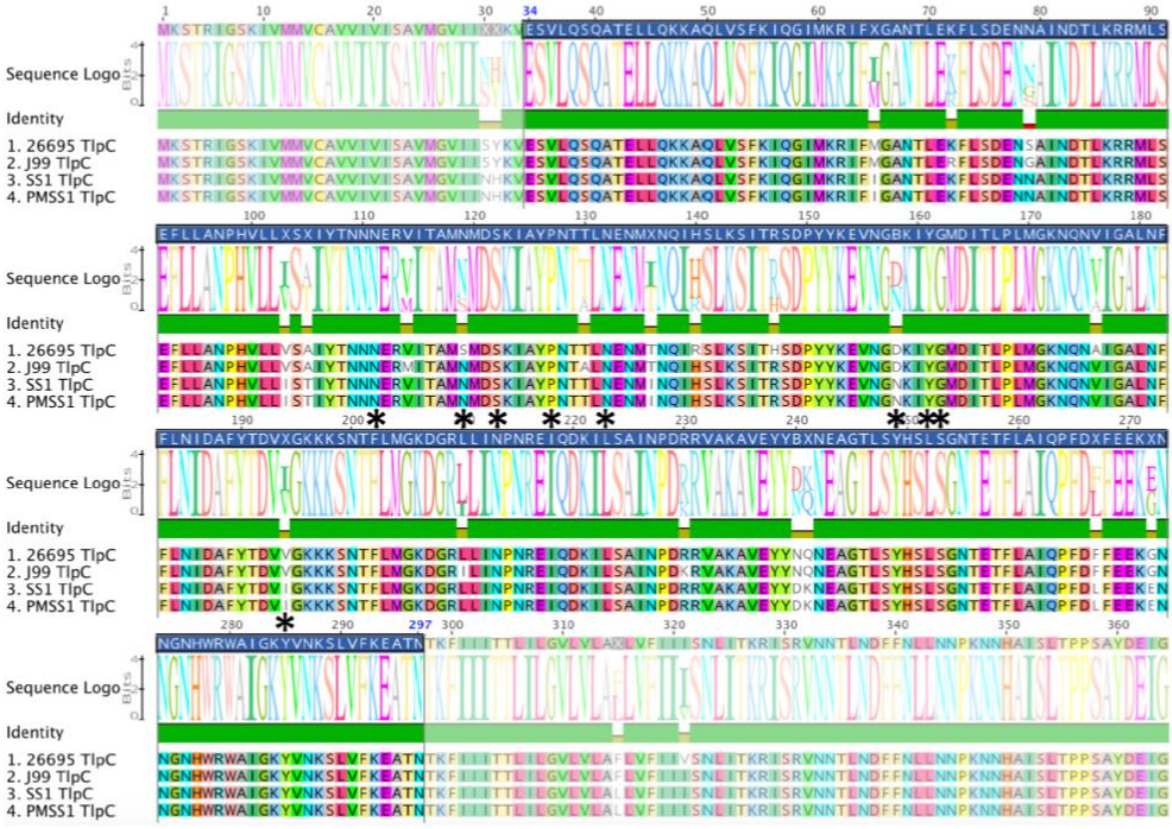
Supplementary information
Supplementary videos 1–12. https://doi.org/10.1038/s41598-017-14372-2.



Supplementary Figure 1. Negative-ion-mode ESI-MS chromatogram. (a) Buffer control, (b) TlpC LBD in same buffer.



Supplementary Figure 2. ITC titrations of TlpC LBD with malate, oxaloacetate, pyruvate and lactate. Upper panel: raw titration data from the injections of 10 µl of 5 mM ligand solution into a 1.45-ml reaction cell containing 10 µM protein. Lower panel: the integrated and dilution corrected peak areas of the titration plot for lactate.



Supplementary Figure 3. Alignment of TlpC protein sequences from H. pylori strains 26695, J99, SS1, and PMSS1. Alignments were produced from whole protein sequences using Geneious version 9.1.8 (Biomatters Ltd) with the Geneious alignment function. Highlighted regions show the ligand binding portion of TlpC that was crystallised. Asterisks indicate the residues of the lactate binding pocket, which are all conserved between 26695 and PMSS1. There is one conservative substitution in the J99 TlpC sequence with L210 being changed to I.

References

1 Schreiber, S. *et al.* The spatial orientation of Helicobacter pylori in the gastric mucus. *Proc Natl Acad Sci U S A* **101**, 5024-5029, doi:10.1073/pnas.0308386101 (2004).

2 Worku, M. L., Karim, Q. N., Spencer, J. & Sidebotham, R. L. Chemotactic response of Helicobacter pylori to human plasma and bile. *Journal of medical microbiology* **53**, 807-811 (2004).

3 Wunder, C. *et al.* Cholesterol glucosylation promotes immune evasion by *Helicobacter pylori*. *Nature medicine* **12**, 1030-1038, doi:10.1038/nm1480 (2006).

4 Cerda, O. A. *et al.* tlpA gene expression is required for arginine and bicarbonate chemotaxis in Helicobacter pylori. *Biological research* **44**, 277-282, doi:/S0716-97602011000300009 (2011).

5 Croxen, M. A., Sisson, G., Melano, R. & Hoffman, P. S. The *Helicobacter pylori* chemotaxis receptor TlpB (HP0103) is required for pH taxis and for colonization of the gastric mucosa. *Journal of bacteriology* **188**, 2656-2665, doi:10.1128/JB.188.7.2656-2665.2006 (2006).

6 Collins, K. D. *et al.* The *Helicobacter pylori* CZB Cytoplasmic Chemoreceptor TlpD Forms an Autonomous Polar Chemotaxis Signaling Complex That Mediates a Tactic Response to Oxidative Stress. *Journal of bacteriology* **198**, 1563-1575, doi:10.1128/JB.00071-16 (2016).

7 Schweinitzer, T. *et al.* Functional characterization and mutagenesis of the proposed behavioral sensor TlpD of *Helicobacter pylori*. *Journal of bacteriology* **190**, 3244-3255, doi:10.1128/JB.01940-07 (2008).

8 Rader, B. A. *et al. Helicobacter pylori* perceives the quorum-sensing molecule AI-2 as a chemorepellent via the chemoreceptor TlpB. *Microbiology* **157**, 2445-2455, doi:10.1099/mic.0.049353-0 (2011).

9 Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244-248, doi:10.1093/nar/gki408 (2005).

10 Machuca, M. A., Liu, Y. C., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni*

Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition. *J. Struct. Biol.* **194**, 205-213, doi:10.1016/j.jsb.2016.02.019 (2016).

11 Liu, Y. C., Machuca, M. A., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. *Acta crystallographica. Section D, Biological crystallography* **71**, 2127-2136, doi:10.1107/S139900471501384X (2015).

12 Zhang, Z. & Hendrickson, W. A. Structural characterization of the predominant family of histidine kinase sensor domains. *Journal of molecular biology* **400**, 335-353, doi:10.1016/j.jmb.2010.04.049 (2010).

13 Machuca, M. A. *et al.* Helicobacter pylori chemoreceptor TlpC mediates chemotaxis to lactate. *Sci Rep* **7**, 14089, doi:10.1038/s41598-017-14372-2 (2017).

Chapter 5: *C. jejuni* Tlp3 dCache sensing domain specifically recognises branched-chain aliphatic amino acids

Preface for Chapter 5

C. jejuni Tlp3 receptor has been shown to directly bind eleven different molecules, including several amino acids: arginine, aspartate, glucosamine, isoleucine and lysine; tricarboxylic acid cycle (TCA) intermediates: malic acid, succinic acid and α-ketoglutarate, as well as purine and thiamine [1], and was proposed as the *Campylobacter* chemoreceptor for multiple ligands (CcmL) [1,2]. The molecules reported to interact with Tlp3 are chemically diverse, thus they are unlikely to be recognised in the exact same mode by the chemoreceptor's LBD. Hence, a combination of biophysical and structural analysis was performed to explore the structural basis of the broad specificity of the Tlp3-LBD. To further explore Tlp3 sensory specificity, a high-throughput screening of its LBD against a library of metabolic substrates was also carried out. This chapter includes a detailed structural analysis of Tlp3-LBD in complex with different amino acids. Methods used and the analysis of the results from this research are presented in the manuscript format which will be submitted to *Cellular and Molecular Life Sciences* (ready for submission).

Monash University

Declaration for Thesis Chapter 5

Declaration by candidate

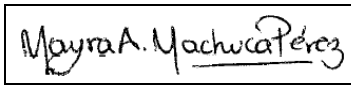
In the case of Chapter 5, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiments, performed experiments, analysed data and wrote manuscript	75

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Mohammad Firoz Khan	Analysed data	**5%**
Anna Roujeinikova	Designed experiments, analysed data, wrote manuscript and led the research	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's Signature		**Date** 10/12/17
Main Supervisor's Signature		**Date** 13 /12/17

Campylobacter jejuni Tlp3 dCache sensing domain specifically recognises branched-chain aliphatic amino acids

Mayra A. Machuca[1,2], Mohammad Firoz Khan[1,2] and Anna Roujeinikova[1,2,3*]

[1]Infection and Immunity Program, Monash Biomedicine Discovery Institute

[2]Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia

[3]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

Abstract

In *Campylobacter jejuni*, chemotaxis and motility have been identified as important virulence factors that are required for host colonisation and invasion. The chemotaxis process involves recognition of chemical cues by the ligand binding domain (LBD) of chemoreceptors also known as methyl-accepting chemotactic proteins. Recently, we determined the crystal structure of *C. jejuni* Tlp3-LBD chemoreceptor in complex with attractant isoleucine, revealing this receptor belongs to the dCache_1 family of sensing modules. In this work, we performed a high-throughput screening of potential ligands and identified additional small molecules that directly interact with Tlp3-LBD. All of the new ligands (leucine, valine and α-amino-N-valeric acid) are amino acids chemically and structurally similar to isoleucine. Analysis of the crystal structures of Tlp3-LBD in complex with these ligands showed that like isoleucine, they bind to the membrane-distal subdomain of the dCache Tlp3 sensing module. The Tlp3-LBD residues that interact with the main chain of isoleucine, leucine, valine and α-amino-N-valeric acid are located at equivalent positions in all complex structures, whilst residues that interact with the side chain move to accommodate the different amino acid ligands.

Introduction

Campylobacter jejuni is a Gram-negative, microaerophilic, flagellated bacterium that colonises the intestines of many wild and domestic animals resulting in a commensal relationship between the bacterium and the host [3-6]. Human infection occurs by direct contact with infected animals or after the consumption of contaminated water, milk or food products [7,8], which may lead to a diarrheal disease. Currently, *C. jejuni* is the leading cause of bacterial gastroenteritis in humans worldwide [9-11]. The clinical manifestations of campylobacteriosis range from an asymptomatic to an acute gastroenteritis characterised by fever, severe abdominal pain and diarrhoea [12]. In most of the cases, *C. jejuni* gastroenteritis is self-limiting, resolving within a week without antibiotic treatment. However, important postinfectious complications of *C. jejuni* infection include reactive arthritis, neuromuscular paralysis, myositis and idiopathic peripheral neuropathy [13,14].

The molecular mechanisms implicated in *C. jejuni* pathogenesis are not well understood. However, chemotaxis has been shown to play an important role in intestinal colonisation in avian and mammalian hosts [15,16]. Non-chemotactic *C. jejuni* mutants, for example, have a reduced capability to colonise the gastrointestinal tract of mice [17,18] and presented a diminished infectivity when compared to their wild-type counterpart [19].

Chemotaxis enables bacteria to sense different cues and find optimal conditions to grow, for example, by moving towards beneficial chemicals (attractants), such as nutrients, or away from harmful compounds (repellents). Chemical cues are sensed by a repertoire of chemoreceptors, also called methyl-accepting chemotaxis proteins (MCPs) or transducer-like proteins (Tlps), which activate a signalling cascade that later controls the rotation of the flagella [20-22]. Most of the studied MCPs are membrane-embedded proteins comprised of an extracytoplasmic ligand binding domain (LBD), a HAMP (histidine kinases, adenylyl cyclases, methyl-accepting protein, and phosphatases) domain and a cytoplasmic methyl-accepting (MA) domain.

Based on their LBDs and membrane topology, chemoreceptors have been classified into seven different classes: I (Ia and Ib), II, III (IIIm and IIIc), and IV (IVa and IVb) [21,23,24]. The class I is a predominant MCP type in both bacteria and archaea species [24]. Chemoreceptors from this class are categorised in two subclasses depending on the number of transmembrane (TM) helices: Ia (two TM helices) or Ib (one TM helix). According to the size of their LBDs, the

subclass Ia MCPs have been further subdivided into two clusters, I (~150 amino acids) and II (~250 amino acids) [24]. Currently, most of the structural information available on chemoreceptor LBDs comes from the study of cluster I receptors, such as the well characterised aspartate receptor Tar from *Salmonella typhimurium* [25] and the serine receptor Tsr from *Escherichia coli* [26], which adopt a 4-helix bundle (4HB) fold. Other characterised cluster I receptor is the urea sensor TlpB LBD from *Helicobacter pylori* that presents a distinctly different fold structure to the 4HB, with a single Cache (calcium channel and chemotaxis) module [27,28]. In cluster II receptors two major structural families (cluster IIa and IIb) have been identified, the helical biomodular (HBM, IIa) [29,30] and double Cache (dCache, IIb) domains [28,31-35].

Analysis of the *C. jejuni* genomes of various different laboratory and clinical strains has identified at least 15 different Tlps or chemoreceptors [20,36-39]. According to their domain organisation, *C. jejuni* Tlps have been categorised into three group, namely as A, B and C [37]. Receptors from group A comprise an N-terminal transmembrane helix followed by a periplasmic LBD, a second transmembrane helix and a C-terminal cytoplasmic MA signalling domain. Members of group A are Tlp1, Tlp2, Tlp3, Tlp4, Tlp7$_{mc}$, Tlp7$_{m}$, Tlp10, Tlp11, Tlp12, and Tlp13. Group B receptor Tlp9 is anchored to the inner membrane that lacks a periplasmic LBD. Group C members Tlp5, Tlp6, Tlp7c and Tlp8, are soluble cytoplasmic proteins that contain a signalling domain homologous to that of group A Tlps.

C. jejuni has been shown to exhibit chemotactic responses to several small molecules, including amino acids, carbohydrates, organic acids, and constituents of bile and mucus (Table 1) [40-42]. However, only a few of these chemical cues have been matched to a specific Tlp (Table 1). For instance, Tlp1 has been shown to detect aspartate [43], whereas Tlp3 is involved in the recognition of multiple ligands, including both attractants, such as isoleucine, purine, aspartic acid, malic acid, fumaric acid and sodium deoxycholate; and repellents including lysine, glucosamine, succinic acid, arginine and thiamine [1,2]. Tlp4 was identified as a sodium deoxycholate receptor [2], Tlp7 as a formic acid receptor [19], and Tlp11 as a galactose receptor [41,44].

Recently, we reported the crystal structure of the LBDs of *C. jejuni* Tlp1 and Tlp3 receptors, both which adopt a dCache_1 fold [32,33]. dCache domains consist of two tandem subdomains (membrane-distal and proximal), with a PAS (Per–Arnt–Sim)-like fold, each of which harbour a putative ligand-binding pocket. These sensory modules have been shown to

directly recognise diverse signal molecules such as amino acids [31,32,45-47], organic compounds [35,47], pyrimidines [48] and purines [49]. In most of the characterised dCache domains, including LBDs of Mlp24 from *Vibrio cholerae* [50], Mlp37 from *V. parahaemolyticus* [47], McpB and McpC from *Bacillus subtilis* [31,51] and Tlp3 from *C. jejuni* [32], direct sensing involved binding of the small signal molecule to the membrane-distal subdomain. However, it has become clear that dCache modules could in principle employ either the membrane-distal and/or membrane-proximal subdomain for ligand recognition [33,35].

For instance, the structural study of the dCache LBD of Tlp1 from *C. jejuni* (performed in Chapter 3) revealed that an acetate ion was bound to the membrane-proximal subdomain [33]. Although acetate has not been suggested to trigger a chemotactic response in *C. jejuni* [40], this finding confirmed that small molecules can also be bound to the membrane-proximal module. Concurringly, the crystallographic and mutational analysis of the dCache LBD of TlpC (described in Chapter 4) showed that the recognition of lactate, TlpC specific ligand, involves direct binding of this small molecule to the membrane-proximal subdomain, instead of the membrane-distal subdomain[35].

Previously in Chapter 2, the crystal structure of the dCache LBD of *C. jejuni* Tlp3 chemoreceptor in free and isoleucine-bound forms (PDB entries 4xmq and 4xmr, respectively), were determined [32]. The structural analysis revealed that isoleucine is recognised through the membrane-distal subdomain of Tlp3-LBD. The Tlp3-LBD/isoleucine complex is stabilised by the formation of multiple hydrogen bonds between the main chain of the amino acid ligand and the side chains of the Tlp3-LBD residues: Lys149, Trp151, Tyr167, Asp169, Asp196 and Thr170. Whilst the aliphatic side chain of the ligand is accommodated in a largely hydrophobic pocket generating van der Waals interactions with the side chains of Tyr118, Val126, Trp151 and Val171 [32]. Noteworthy, five of the residues (Lys149, Trp151, Tyr167, Asp169 and Asp196) interacting with the amino and carboxyl groups of the ligand, were found to be strongly conserved in dCache sensory modules that recognise amino acid ligands [32]. In contrast, the residues that only interact with the side chain of the ligand are not conserved and instead, substitutions in these residues have been suggested to determine the specificity of the amino acid receptor [32]. These findings provided evidence of the mechanism of recognition of the main chain of an amino acid, and suggest that Tlp3-LBD is likely to bind other similar amino acids to its membrane-distal subdomain.

Table 1. *C. jejuni* chemotactic behaviour.

Compound	Chemotactic response[a]	Chemoreceptor matched	Reference
Carbohydrates			
Galactose	+	Tlp11	41
L-fucose	+	?	40,52
Amino acids			
L-Arginine	-	Tlp3	1
L-Aspartate	+	Tlp1, Tlp3	1,40,42,43
L-Asparagine	+	?	19,42
L-Cysteine	+	?	19,40,42
L-Glutamate	+	?	19,40,42
L-Glucosamine	-	Tlp3	1
L-Isoleucine	+	Tlp3	1,32
L-Lysine	-	Tlp3	1
L-Serine	+	?	19,40,42
Organic acids			
D-Lactate	+	?	19,42
L-malate (malic acid)	+	Tlp3	1,19,42
Formic acid (Formate)	+	Tlp7	19,42
Fumarate	+	?	19,42
α-Ketoglutamate	+	?	42
α-ketoglutarate	+	Tlp3	1
Pyruvate	+	?	19,42
Succinate	+ and -	Tlp3	1,42
Mucin constituents			
Galactose	+	Tlp11	41
L-fucose	+	?	40,52
Bile constituents			
L-fucose	+	?	40,52
Cholic acid	-	?	19,40,42
Deoxycholic acid	-	?	19,40,42
Sodium deoxycholate	+	Tlp3, Tlp4	2
Glycocholic acid	-	?	19,40,42
Glycochenodeoxycholic acid	-	?	40
Glycodeoxycholic acid	-	?	40
Taurocholic acid	-	?	19,40,42
Other compounds			
Purine	+	Tlp3	1
Thiamine	-	Tlp3	1

[a]**Positive chemotactic response (+) and negative chemotactic response (-)**

Besides isoleucine, Tlp3-LBD has been shown to directly bind other ten molecules and was proposed the *Campylobacter* chemoreceptor for multiple ligands (CcmL) [1,2]. The reported molecules that interact with Tlp3 are chemically diverse. Thus they are unlikely to be recognised in the same mode by the chemoreceptor's LBD. To explore the structural basis of the broad specificity of the Tlp3-LBD, a high-throughput screening of this LBD against a previously used library of metabolic substrates for characterising other LBD chemoreceptors was performed [45,53,54]. Screening and microcalorimetric assays included ten ligands that have been reported that directly bind to Tlp3-LBD. Tlp3-LBD was found to bind only three ligands which are chemically similar to isoleucine (L-leucine, L-valine and α-amino-N-valeric acid). The crystal structures of the Tlp3-LBD complex with L-leucine, L-valine and α-amino-N-valeric acid were determined to a resolution of 1.4 Å, 1.4 Å and 2.1 Å, respectively. Analysis of Tlp3-LBD/complex structures showed that these ligands are bound in a very similar fashion at the membrane-distal subdomain and revealed the structural changes of the ligand-binding pocket which allows Tlp3-LBD to accommodate these three ligands.

Methods

Thermal shift assays

The LBD of Tlp3 chemoreceptor from *C. jejuni* serotype O:2 (strain NCTC 11168, residues 42–291 plus GIDPFT sequence at the N terminus as a cloning artefact) was expressed in *E. coli* using the pET151/D-TOPO vector and purified following the previously published procedure [55]. A high-throughput screening of metabolic substrates was performed by thermal shift assays following a slightly modified protocol by McKellar *et al.* [45]. In these assays, a mixture of the purified LBD and a fluorescent compound is exposed to a temperature gradient; and the protein unfolding curve is monitored by detecting changes in the fluorescence. Thermal melt assays permit the calculation of the melting temperature (T_m) value, which corresponds to the temperature at which 50% of the protein has denatured and is a measure of the thermal stability of the protein [56]. Ligand binding to a protein generally alters the thermal stability of the protein, which can be evinced by change in the T_m value [57,58].

The compounds from the Biolog Phenotype Microarray (PM) PM1, PM3B and PM5 plates were used for the high-throughput screening. These plates comprise a library of potential metabolic substrates including carbon and nitrogen sources, and nutrients

supplements [59]. Each plate contains 95 compounds and a negative control (water only) (Supplementary Table S1). Ligands were prepared by dissolving the Biolog compounds in 50 µl of water to a final concentration of 10-20 mM (as indicated by the manufacturer). Each assay mixture was prepared to a final volume of 25 µl and contained 20 µM protein, 10 mM Tris-HCl pH 8.0, 150 mM NaCl and SYPRO Orange (SIGMA) at 5x concentration. Two microliters of the resuspended Biolog compounds were added to each well. The samples were heated from 35°C to 80°C at a ramp rate of 0.5°C min^{-1} using a Rotor-Gene Q real-time PCR cycler (Qiagen, Venlo, Netherlands). All the measurements were performed in triplicates. The T_m of ligand-free protein or in the presence of each compound were calculated using nonlinear regression (curve fit) implemented in the GraphPad Prism software (version 7.02).

Isothermal titration calorimetry (ITC)

His6-tagged Tlp3 LBD was dialysed against 100 mM Tris-HCl pH 8.0, 150 mM NaCl. Solutions of 3 mM L-leucine and α-amino-N-valeric acid, and 10 mM L-valine, 2'-deoxycytidine, Ala-Thr, methyl pyruvate, L-lysine, L-arginine, L-aspartate and purine were prepared by dissolving them in the dialysis buffer. Experiments were conducted on a VP-ITC MicroCal calorimeter (Malvern Instruments, UK) at 25°C. The protein sample at 10 µM concentration in a 1.45 ml reaction cell was titrated with 25 successive 10 µl injections of ligand solution at a spacing of 300 s. All measurements were made in triplicate. The data obtained were integrated and normalised for protein concentration to generate binding isotherms. These were corrected with heat changes measured for control experiments where the same concentration of ligand was titrated into the buffer. The baseline-corrected data were fitted to an equation for the "one binding site" model using Origin ITC software (OriginLab, USA), fixing the stoichiometry (N) as 1 and floating all other fitting parameters to calculate the dissociation constant (K_D).

Protein crystallisation, data collection and structure determination

To co-crystallise Tlp3-LBD (residues 42–291 plus GIDPFT sequence at the N terminus as a cloning artefact) with L-leucine, L-valine, α-amino-N-valeric acid, L-lysine, L-aspartate, L-arginine and malic acid; the protein was concentrated to 15 mg ml^{-1} in a buffer containing 10 mM Tris-HCl pH 8.0, 150 mM NaCl and incubated with 8 mM of each compound for 30 min. The solution was then clarified by centrifugation at 13,000 g, 4°C for 20 min and protein crystallisation was carried out by the hanging-drop vapour diffusion method using the crystallisation conditions previously described [55], with a few modifications. Briefly, crystals

were grown in 2 µl hanging drops suspended over 500 µl reservoir solution containing 22% (w/v) polyethylene glycol 3350, 100 mM sodium citrate pH 5.0 and 200 mM ammonium sulphate, where 1 µl protein solution was mixed with an equivalent aliquot of the reservoir solution. Crystals of the Tlp3-LBD complexes were isomorphous and belonged to the space group $P2_1$ (Table 2), with two molecules in the asymmetric unit.

To perform data collection at cryogenic temperatures, Tlp3-LBD crystals were washed for 1 s in a cryoprotectant solution consisting of 26% (w/v) polyethylene glycol 3350, 100 mM sodium citrate pH 5.0, 200 mM ammonium sulphate, 10% (v/v) glycerol, 10 mM L-leucine, L-valine, α-amino-N-valeric, L-lysine, L-aspartate, L-arginine or malic acid, and were flash-cooled by plunging them into liquid nitrogen. X-ray diffraction data were collected at -173.2°C on the MX1 and MX2 beamlines of the Australian Synchrotron (AS) [60] to 1.4 Å, 1.4 Å and 2.1 Å resolution for Tlp3-LBD/leucine, /valine and /α-amino-N-valeric acid complex, respectively. X-ray diffraction data were also collected for Tlp3-LBD crystals co-crystallised with L-lysine, L-arginine, L-aspartate and malic acid to a 1.4, 1.45, 1.5 and 1.45 Å resolution, respectively. All data were processed using iMOSFLM [61] and scaled with AIMLESS [62] from the CCP4 software suite [63]. A summary of the data processing statistics for Tlp3-LBD/leucine, /valine / and /α-amino-N-valeric acid complex is presented in Table 2.

The Tlp3-LBD/co-crystallised structures were determined using molecular replacement with PHASER [64] and the coordinates of Tlp3-LBD/isoleucine complex (PDB code, 4xmr) [32] as a search model. The models were refined with PHENIX [65] and manually rebuilt where necessary using COOT [66]. The final models were validated using MOLPROBITY [67] and the refinement statistics for each model are summarised in Table 3. The structure figures were produced using PYMOL 4.1.

Table 2. X-ray data collection and processing statistics. Values in parentheses are for the highest resolution shell.

Dataset	Co-crystal with 8 mM leucine	Co-crystal with 8 mM valine	Co-crystal with 8 mM α-amino-N-valeric acid
Space group	$P2_1$	$P2_1$	$P2_1$
a, *b*, *c* (Å)	42.6, 137.5, 48.9	42.6, 137.5, 49.1	41.9, 137.8, 49.2
β (°)	94.3	94.5	93.9
Resolution range (Å)	28.21-1.42 (1.42-1.40)	28.18-1.42 (1.42-1.40)	46.23-2.16 (2.16-2.10)
R_{merge}	0.048 (0.399)	0.047 (0.308)	0.077 (0.313)
Average I/σ(I)	10.7 (2.6)	11.4 (3.1)	10.6 (3.3)
Completeness (%)	99.2 (99.4)	94.3 (88.0)	97.6 (96.6)
Redundancy	3.7	3.2	4.3
Observed reflections	403,905	328,682	134,763
Unique reflections	109,203	103,884	31,594

Table 3. Refinement statistics

Dataset	Co-crystal with leucine	Co-crystal with valine	Co-crystal with α-amino-N-valeric acid
Resolution range (Å)	28.2-1.4	28.2-1.4	39.9-2.1
No. reflections	108,818	103,732	31,561
R_{work}/ R_{free} [a]	0.168/0.189	0.169/0.191	0.168/0.214
No. atoms			
Protein	4465	4489	4221
Water	820	944	241
B-factors			
Average *B* (protein atoms) ($Å^2$)	15	14	26
Average *B* (water molecules) ($Å^2$)	29	29	30
Average *B* (ligand) ($Å^2$)	10	10	15
r.m.s. deviations from ideality			
Bond lengths (Å)	0.005	0.005	0.006
Bond angles (º)	0.768	1.797	0.807

[a] The R_{free} was calculated on 5% of the data omitted at random.

Results

Identification potential ligands of Tlp3 by thermal shift assays

To investigate the ligand-binding profile of Tlp3-LBD, a high-throughput thermal shift screening of a library of metabolic substrates that may potentially interact with this chemoreceptor's LBD was performed. Tlp3-LBD was screened against 248 different compounds from the Biolog PM plates PM1, PM3 and PM5, which include different carbon and nitrogen sources and nutrients supplements [59]. Although PM plates are designed to characterise the metabolic capabilities of microbial species, it was previously shown that they can be used in thermal shift assays as a library of potential ligands to identify the specificity of chemoreceptors [45,49,54].

The midpoint of unfolding (T_m) of free Tlp3-LBD was 56.3 ± 0.1°C. The ability of each of the compounds tested to alter the thermal stability of the protein was quantified by its thermal shift (ΔT_m): as the Tm in the presence of the compound minus the Tm in the absence of the compound (ligand-free protein). The ΔT_m induced by compounds of the PM1, PM3 and PM5 plate are shown in Figure 1. A total of six compounds were identified to increase the Tm value by at least 1°C (Figure 1A, Table 4, Table S1), which was considered as the threshold for a potential compound-Tlp3-LBD interaction in this screening. Our analysis showed that L-leucine and L-isoleucine caused the most pronounced Tm increase with a ΔT_m of 2.4°C and 2.1°C, respectively; followed by 2-deoxy cytidine and Ala-Thr that induced an increase of 1.4 and 1.1°C in the T_m, respectively. We also observed that many of the molecules tested produced a decrease of the T_m by -0.5 to -1.8°C (Figure 1A). Among these, methyl pyruvate was the compound that caused the most significant decrease in protein T_m (ΔT_m of -1.8 °C, Figure 1A). Therefore, methyl pyruvate was also considered as a hit from the screening.

The library of metabolic substrates screened includes nine of the previously reported molecules by Rahman *et al.* [1] that bind to Tlp3-LBD. The ΔT_m induced by each of these nine molecules are summarised in Table 4. Surprisingly, none of them had a significant effect on protein T_m and were not considered as hits from the high-throughput screening.

Confirmation of ligand binding to Tlp3-LBD by ITC

ITC experiments were carried out with L-leucine, L-valine, α-amino-N-valeric acid, methyl pyruvate, Ala-Thr and 2-deoxy cytidine; and with isoleucine as a positive control. Additionally, microcalorimetric experiments were also performed with four of the previously

reported molecules that bind to Tlp3-LBD by Rahman *et al.* (L-lysine, L-arginine, L-aspartate and purine) [1]. The determined thermodynamic parameters are summarised in Table 4

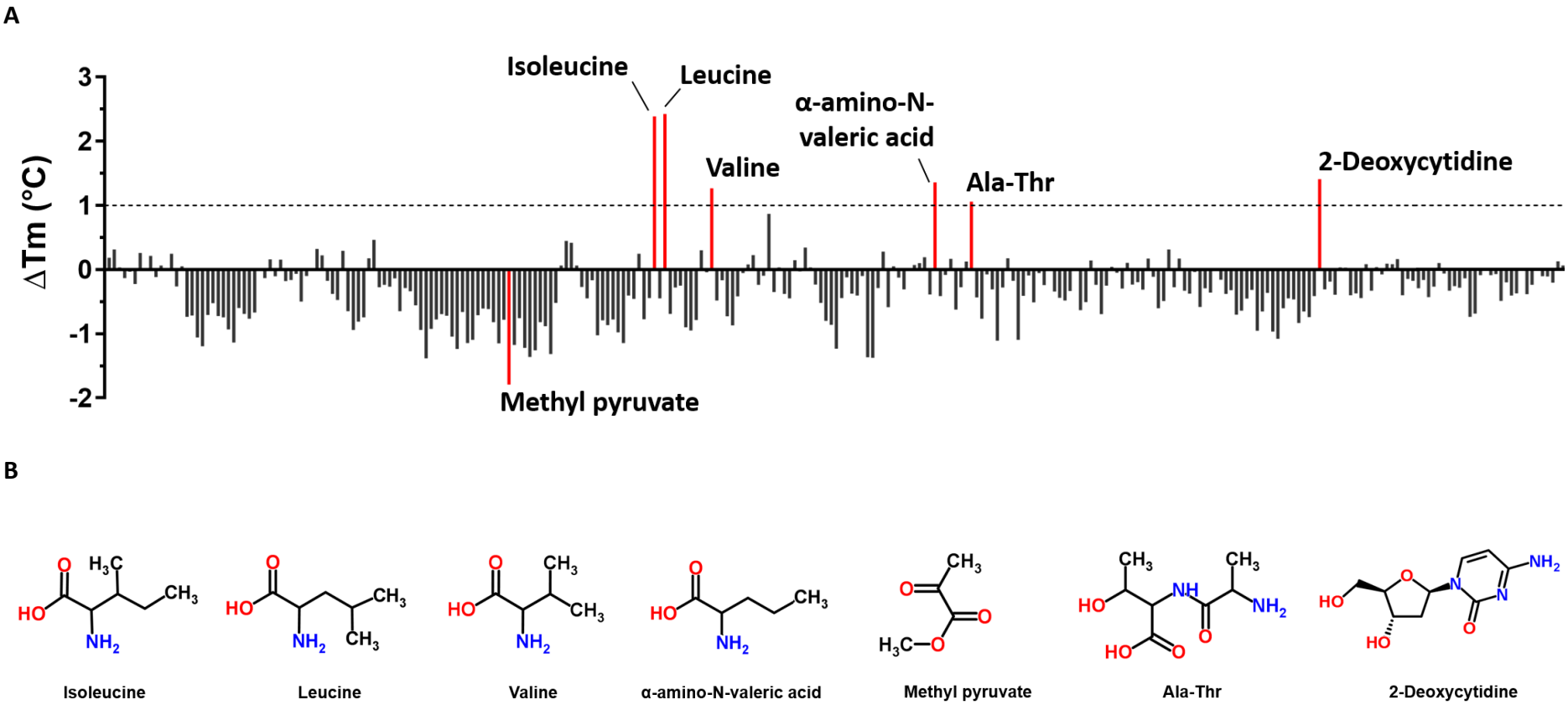


Figure 1. Results of thermal shift assays performed on (A) Tlp3-LBD in the presence of the compounds from the Biolog screens PM1, PM3B and PM5. The ΔT_m was calculated as the T_m of the protein in the presence of compound minus the T_m in the absence of the compound (56.3 °C). The full list of all tested compounds and the respective ΔT_m is provided in supplementary table S1. Compounds that caused T_m shifts of at least 1°C are labelled. (B) Structures of the compounds that were identified as hits in the thermal shift assay. The chemical structures were downloaded from ChemSpider database (http://www.chemspider.com/About.aspx) [68].

The ITC measurements performed revealed that from the ten compounds tested Tlp3-LBD only bind the amino acids L-leucine and L-valine; and the non-common amino acid α-amino-N-valeric acid (an isomer of valine). Titration of the protein with these three ligands caused exothermic heat changes which confirmed their binding to Tlp3-LBD (Figure 2, Table 4). Data analysis using a model for single binding site revealed a KD of 104.5 µM, 404.8 µM, 167.5 µM for leucine, valine and α-amino-N-valeric acid, respectively; with the binding process driven by favourable enthalpy changes for the three ligands (Figure 2A-C, Table 4). No binding between Tlp3-LBD and L-lysine, L-arginine, L-aspartate nor purine could be detected by ITC experiments (Figure 2D, Table 4).

Table 4. Thermal shift and ITC measurements data

Compound	Hits[b]	Thermodynamic parameters[c]		
	ΔT_m (°C)	**K_D (µM)**	**ΔH (kcal mol^{-1})**	**ΔS (cal/mol/deg)**
L-Ile[a]	2.1 ± 0.8	86.2 ± 10.2	-4.4 ± 0.2	3.6
L-Leu	2.4 ± 0.2	104.5 ± 5.9	-4.6 ± 0.1	2.9
L-Val	1.3 ± 0.1	404.8 ± 26.5	-4.4 ± 0.2	0.8
α-amino-N-valeric acid	1.4 ± 0.3	167.5 ± 8.7	-6.5 ± 0.2	-4.5
Ala-Thr	1.1 ± 0.3	NB[d]		
2-deoxycytidine	1.4 ± 0.8	NB		
Methyl pyruvate	-1.8 ± 0.0	NB		
L-lysine	-0.5 +- 0.1	NB		
L-arginine	0.1 ± 0.2	NB		
L-aspartic acid	-0.2 ± 0.7	NB		
D-glucosaminic acid	-0.4 ± 0.3	NP[e]		
Succinic acid	-0.1 ± 0.6	NP		
Malic acid	-0.8 ± 0.4	NP		
Thiamine	-0.1 ± 0.4	NP		
α-ketoglutaric acid	-0.1 ± 0.0	NP		
Purine	NP	NB		

[a] **K_D for L-isoleucine was determined by Liu *et al.* [32].**

[b] **Hits identified by thermal shift assays.**

[c] **Determined by ITC measurements.**

[d] **NB: No binding detected by ITC measurements**

[e] **NP: Not performed**

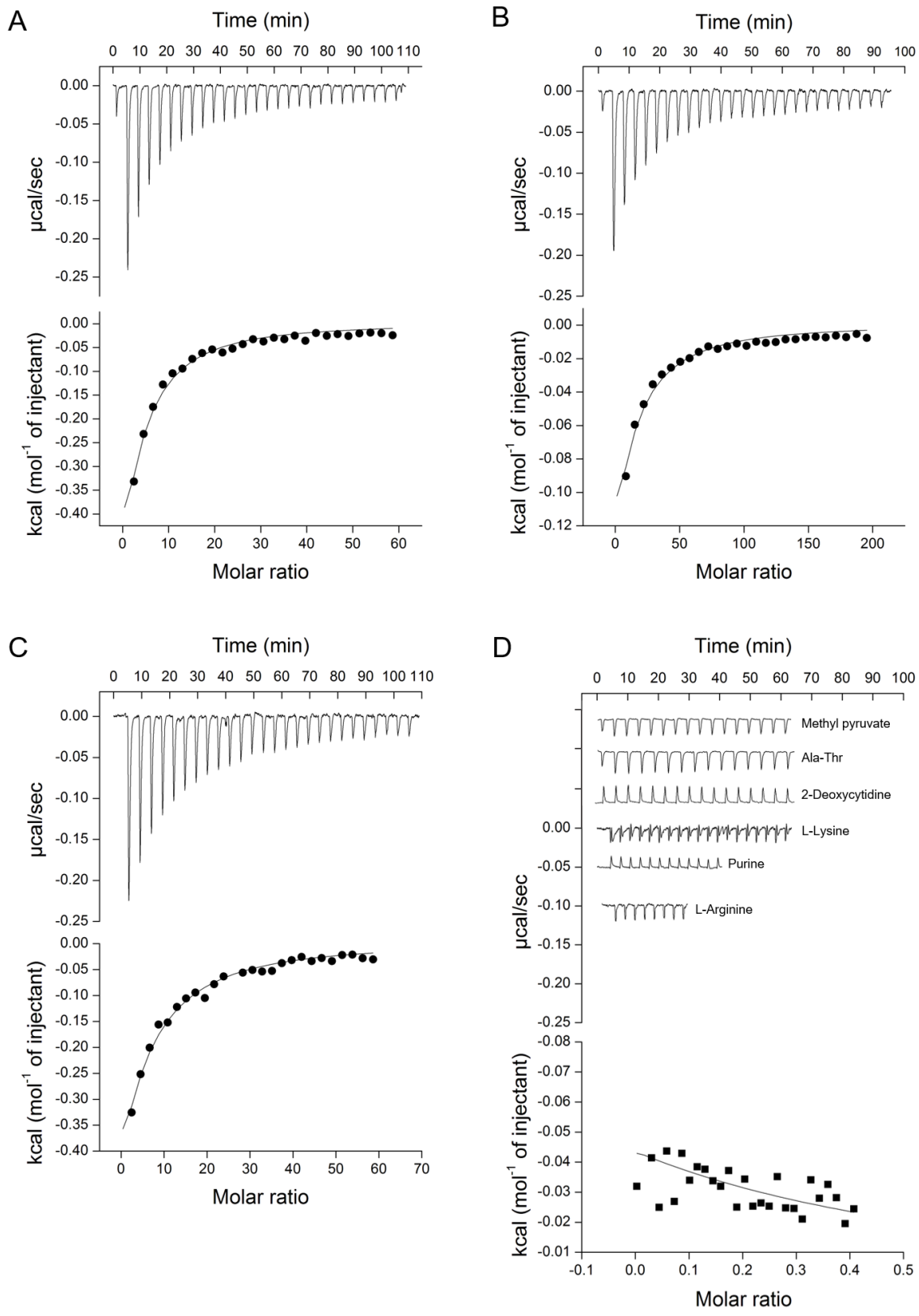


Figure 2. Tlp3-LBD binds the amino acids leucine, valine and α-amino-N-valeric acid. Representative ITC data of Tlp3-LBD titrated with (A) L-leucine, (B) L-valine, (C) α-amino-N-valeric acid, (D) Ala-Thr, 2-deoxycytidine, L-lysine, purine or L-arginine. The upper panels

show raw titration data. Lower panels are the integrated and dilution corrected peak areas of the titration data. For the titrations, Tlp3-LBD was at 10 µM, and the ligands were at 3 mM for leucine and α-amino-N-valeric acid; and 10 mM for valine or 10 Mm for methyl pyruvate, Ala-Thr, 2-deoxycytidine, L-lysine, purine or L-arginine.

3D structure of Tlp3-LBD in complex with L-leucine, L-valine and α-amino-valeric-acid

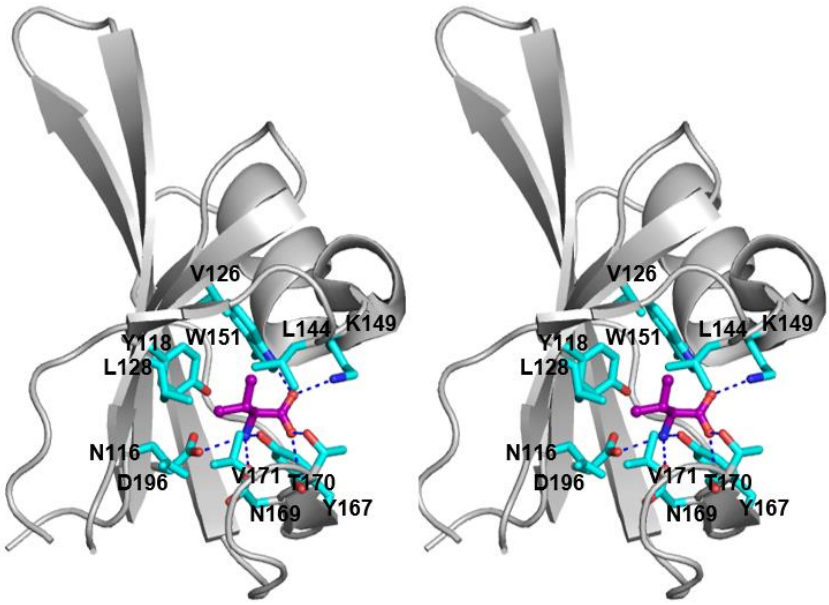
To investigate the structural basis for ligand recognition in the dCache Tlp3-LBD, the crystal structures of Tlp3-LBD in complex with the ligands identified in the screening and confirmed by ITC L-leucine, L-valine and α-amino-N-valeric acid, were determined. The crystal structures of Tlp3-LBD/leucine, /valine and /α-amino-N-valeric acid complexes belonged to the space group $P2_1$, with two molecules in the asymmetric unit related to each other by two-fold axis symmetry, as previously observed for the Tlp3-LBD free form and isoleucine complex structures analysed in detail in Chapter 2 (4xmq and 4xmr, respectively) [32]. Structures of Tlp3-LBD co-crystallised with 8 mM L-lysine, L-arginine, L-aspartate or malic acid were also determined to further confirm that these molecules do not directly interact with Tlp3 dCache LBD. Analysis of the pseudo-atomic resolution (1.4, 1.45, 1.5 and 1.45 Å, for L-lysine, L-arginine, L-aspartate and malic acid; respectively) Fourier maps revealed no electron density that could be interpreted as any of these molecules bound to Tlp3-LBD, which is in agreement with the results obtained from the binding assays performed.

Analysis of Fourier difference maps of the Tlp3-LBD/leucine, /valine and /α-amino-N-valeric acid complex structures revealed unambiguous electron density corresponding to the ligands bound at the membrane distal pocket of the dCache LBD. The amino acid atomic backbone atoms of the three ligands are located in an almost identical position to that of isoleucine in Tlp3-LBD/isoleucine complex structure [32]. Calculation of the accessible surface area (ASA) showed that the ligands are completely shielded from the solvent upon binding to Tlp3-LBD, with > 99% of their ASA buried by the protein (buried surface area of 99.4%, 99.6% and 99.5% for L-leucine, L-valine and α-amino-N-valeric acid; respectively)

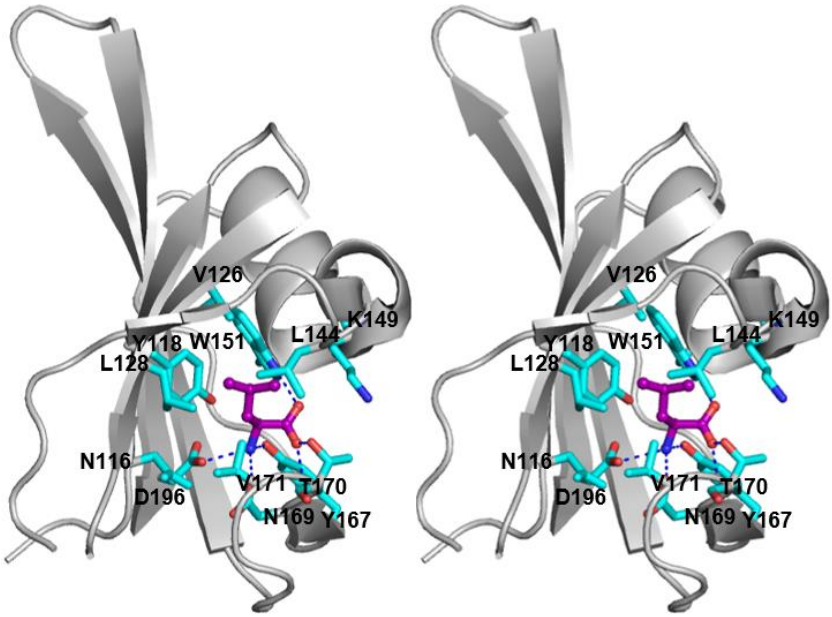
In the Tlp3-LBD/valine complex, the amino acid ligand is stabilised by the formation of seven hydrogen bonds, which are mostly conserved in the other Tlp3 ligand-bound structures, including the isoleucine complex (Figure 3) [32]. The L-valine amino group is forming hydrogen bonds with the Tyr167 O^{η}, Asp169 $O^{\delta 1}$, and Asp196 $O^{\delta 2}$; while the carbonyl oxygen interacts with the Thr170 $O^{\gamma 1}$ and the main-chain amino group of the same residue. Furthermore, the

OXT group forms hydrogen bonds with the Lys149 N^{ζ} and Trp151 $N^{\varepsilon 1}$ and the aliphatic side chain of L-valine is making van der Waal contacts with the side chain of Tyr118, Leu144, Trp151 and Val171, and is 4.2 Å from the Asn116 C^{β} and Leu128 $C^{\delta 1}$.

A



B



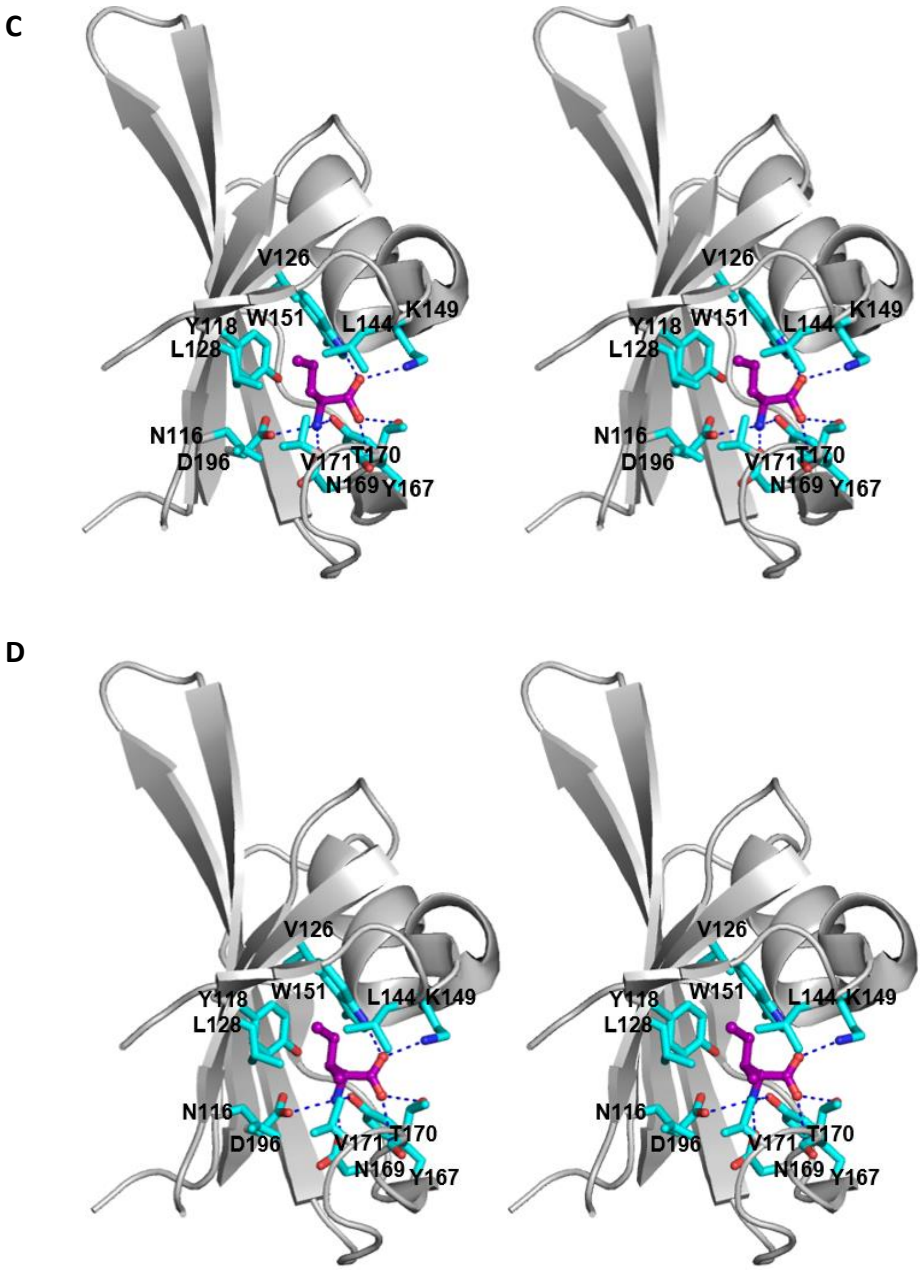


Figure 3. Close up view of the ligand-binding pocket at the membrane-distal subdomain of the Tlp3-LBD/valine (A), /leucine (B), /α-amino-N-valeric acid (C) and /isoleucine (D) complex. The ligand molecules are shown in all-atom ball-and-stick representation with C atoms coloured purple. The protein side chains that form direct contacts with the ligands are shown in stick representation.

Analysis of the binding mode of L-leucine and α-amino-N-valeric acid showed that both ligands form similar interactions to those observed in the L-valine and L-isoleucine complex. In the L-leucine and α-amino-N-valeric acid structures, their amino and carboxyl groups are

located at equivalent positions to that of L-valine and interact with the Tlp3 residues Trp151, Tyr167, Thr170, Asp169 and Asp196. The binding of α-amino-N-valeric acid and L-leucine to Tlp3-LBD is also favoured by the formation of van der Waals contacts with the residues Tyr118, Trp151, Val 126 and Val171. Furthermore, the side chain of leucine is approaching within 4.1 and 4.2 Å of the side chains of L144 and L128; respectively, while the α-amino-N-valeric acid side chain is within 3.9 and 4.3 Å of L144 and L128; respectively.

The overall structural features of L-leucine, L-valine and α-amino-N-valeric acid Tlp3-LBD complexes are very similar to each other and similar to that of L-isoleucine complex (Figure 4). Superimposition of the three structures showed no significant differences, with an average root-mean-square deviation (r.m.s.d.) value of 0.23 Å. Analysis of the molecular surface of Tlp3-LBD/ complex structures using CASTp [69] with a probe radius of 1.4 Å revealed a change in the volume of the putative binding pocket at the membrane-distal subdomain (Figure 5). The solvent-accessible volume of the membrane distal pocket in Tlp3-LBD/valine complex structure is calculated to be 207.2 $Å^3$ and increases to 235.7 $Å^3$ and 240 $Å^3$ in the /leucine and /isoleucine complex structures; respectively, which is proportional to the volume calculated for each amino acid ligand in water (Figure 5).

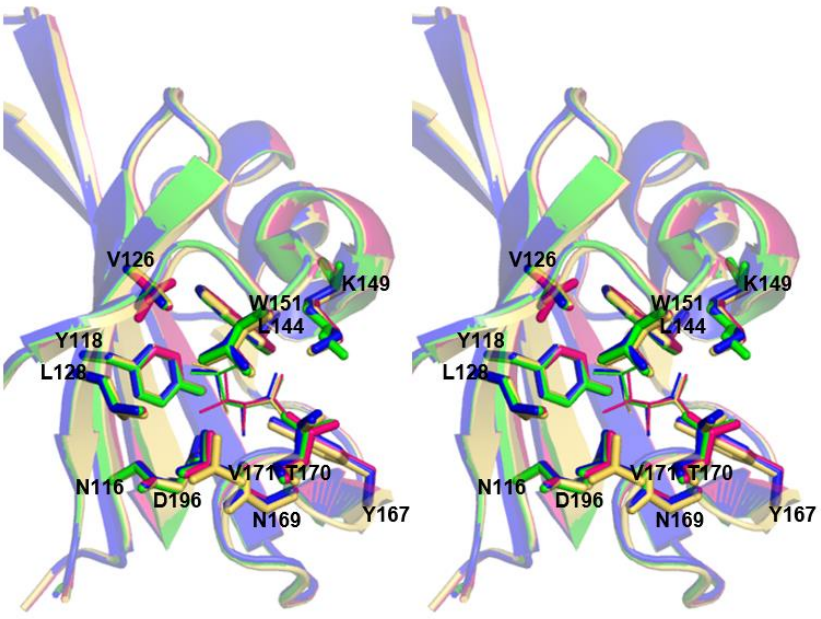


Figure 4. Comparison of the membrane-distal subdomain of Tlp3 dCache LBD in complex with different ligands. Superposition of the of Tlp3 LBD/valine complex (pink) with Tlp3 LBD/leucine (green), Tlp3 LBD/α-amino-N-valeric acid (blue) and Tlp3 LBD/isoleucine (yellow)

complexes. The residues involved in ligand recognition are shown as sticks, and ligand molecules are presented as lines.

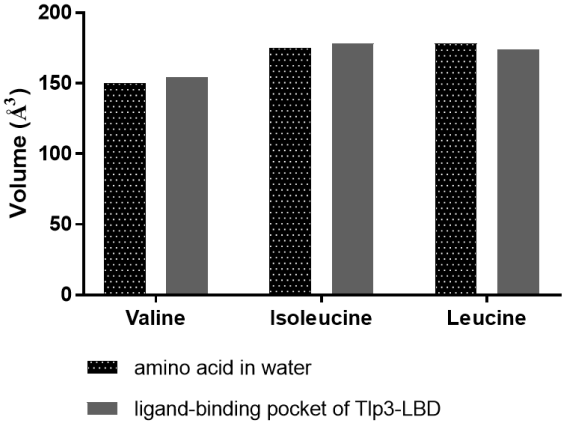


Figure 5. Changes in the probe-accessible volume of the ligand-binding pocket in the membrane-distal subdomain induced by binding of ligands. For calculations atoms of the ligands have been excluded from the model.

To further analyse the conformational changes the membrane distal pocket undergoes to accommodate different amino acid ligands, the membrane-distal subdomain (residues 63-197) of the four Tlp3-LBD/ complex structures were superimposed and analysed for Cα atom r.m.s.d. (Figure 6). The most significant differences are observed in the structures of the loops β2'α3 and β3β4 tongue which form part of the binding pocket (Figure 6A), with the following residues showing the largest displacements of Cα atoms: Asp143 (0.53 Å), Leu144 (0.42 Å), Lys147 (0.81 Å), Thr148 (1.17 Å), Lys149 (0.46 Å) (loop β2'α3), Val171 (0.82 Å), Leu172 (0.43 Å) and Lys173 (0.41 Å) (loop β3β4). Analysis of the distribution of the main-chain temperature factor averaged over the four Tlp3-LBD/ complex structures also indicates these two loops as variable regions in the structure (Figure 6B).

Additionally, structure superimposition shows that the side chain of some residues in the two variable loops have a different conformation. For instance, in the Tlp3-LBD/valine complex the side chain of Leu144 (loop β2'α3) comes closer to the L-valine C^{β} and C^{δ} atoms. Meanwhile, in the other complex structures Leu144 side chain moves to accommodate the $C^{\delta 1}$ atom of isoleucine or α-amino-N-valeric acid; or $C^{\delta 1}$ and $C^{\delta 2}$ atoms of leucine. Similarly, the side chain of Val171 (loop β3β4) in the Tlp3-LBD/isoleucine complex is positioned 1.5 Å

away from the binding pocket core in comparison to its spatial location in the other complexes, which allows accommodating the $C^{\gamma 2}$ atom of isoleucine.

The Tlp3-LBD residues (Lys149, Trp151, Tyr167, Asp169 and Asp196) that interact with the amino and carboxyl groups of the ligands are located at equivalent positions in all the complex structures. These residues were previously described in Chapter 2 to be highly conserved in different dCache amino acid receptors. Interestingly, the conserved residue Lys149 is forming a hydrogen bond with the OXT group of the amino acid ligands, except for L-leucine. This interaction is disrupted in the Tlp3-LBD/leucine complex, where the side chain of Lys149 change into a different rotamer conformation, with the N^{ζ} atom facing away from the ligand-binding pocket (Figure 7) to accommodate the $C^{\delta 2}$ atom of leucine.

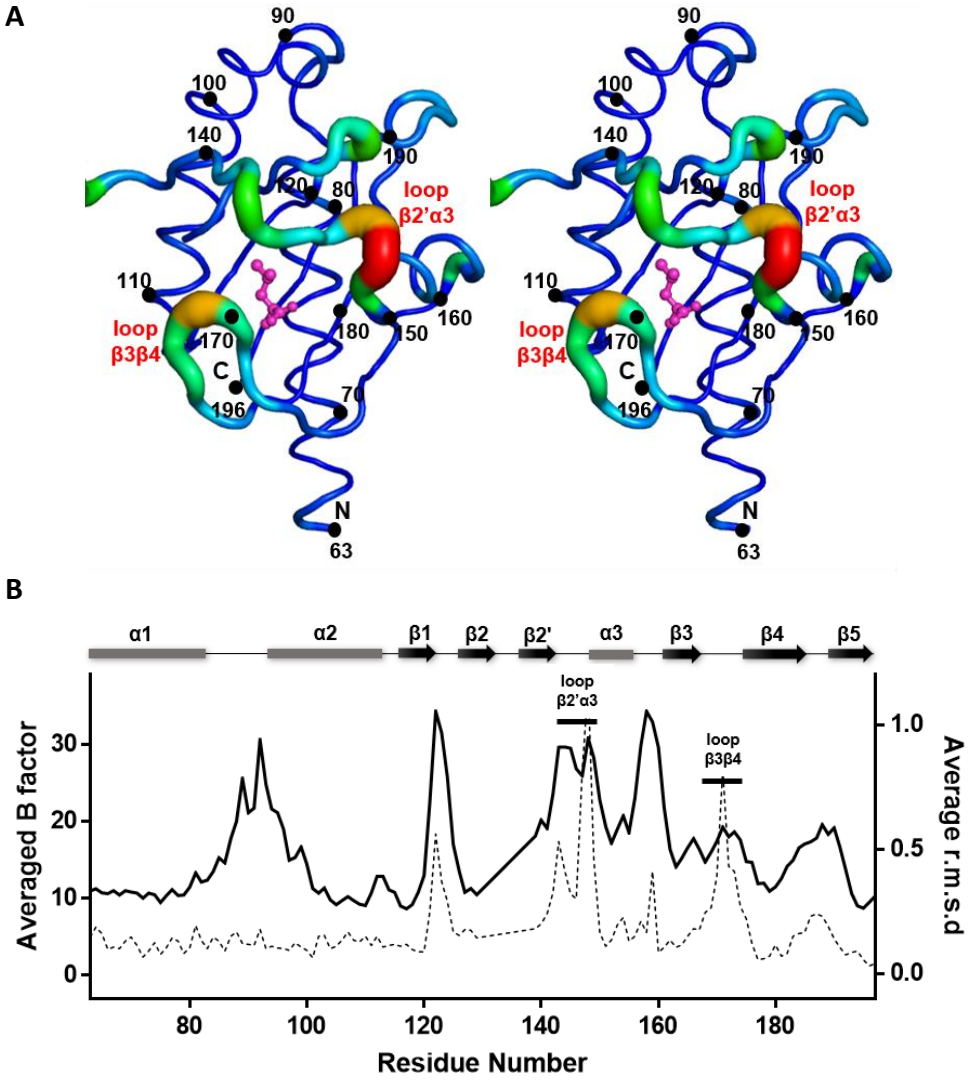


Figure 6. Comparison of the membrane-distal subdomain of Tlp3-LBD/ complex structures. (A) Stereo diagram of the structure of the membrane-distal subdomain of dCache Tlp3

receptor. The backbone radius is proportional to the average Cα atom r.m.s.d. to the mean structure for the superimposition of the subunit B membrane-distal subdomain from all four Tlp3-LBD/complexes. r.m.s.d. values were calculated using Superpose from the CCP4 suite. The colour gradient runs from blue (the smallest r.m.s.d.) to red (the largest r.m.s.d.). (B) Average main chain temperature B factor as a function of residue number. The crystallographic B values have been averaged from the subunit B of the four Tlp3-LBD/ complex structures (solid line). Average Cα atom r.m.s.d. is shown as dotted line. The positions of the two variable loops are indicated.

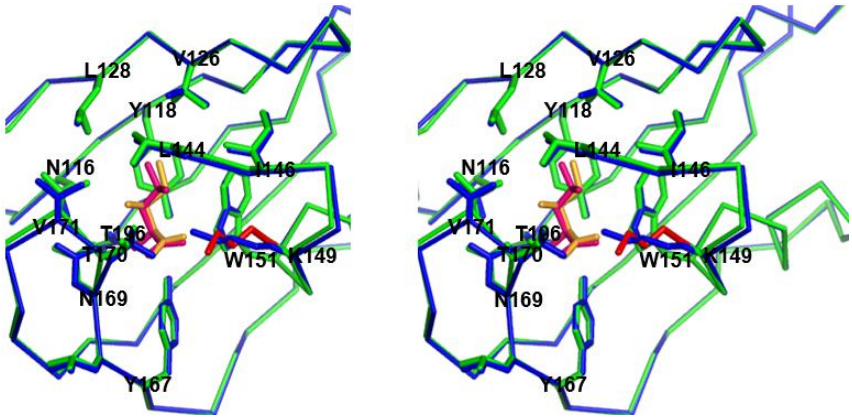


Figure 7. Comparison of the membrane-distal subdomain of dCache Tlp3-LBD/leucine and /isoleucine complex structures. Superposition of the Tlp3 LBD/leucine (green) and Tlp3 LBD/isoleucine (blue) complexes. Stereo diagram highlighting the conformational change of the residue K149 (shown in red) in the Tlp3 LBD/leucine complex. The residues involved in ligand recognition are shown as sticks. The leucine (shown in pink) and isoleucine (shown in orange) ligand molecules are presented as sticks.

Conformational changes after ligand binding to Tlp3-LBD

Previously in Chapter 2, the structural analysis of the apo Tlp3-LBD (ligand-free) and Tlp3-LBD/isoleucine complex structures was described. In the previous analysis, superimposition of the two Tlp3-LBD molecules present in the asymmetric unit showed that in one subunit the membrane-proximal subdomain adopts a more open form than the other, with a downward ~4 Å displacement of the C-terminal helix. The difference in conformation was observed in both the apo Tlp3-LBD and the isoleucine complex. Thus, the structures of subunit A in the apo Tlp3-LBD structure, which presents the membrane-distal subdomain in an open conformation and membrane-proximal in a closed; and subunit B in the isoleucine

complex, with the membrane-distal subdomain close and membrane-proximal open, represent the extreme conformational states of free and ligand-bound Tlp3-LBD. Comparison of these subunits revealed significant conformational changes that Tlp3-LBD underwent after isoleucine was bound [32].

To determine whether the changes identified in the Tlp3-LBD/isoleucine structure also occur in the other complexes, superimposition of the subunit B of Tlp3-LBD/valine, /leucine /α-amino-N-valeric acid complex and the subunit A of apo Tlp3-LBD was performed. As observed in the Tlp3-LBD/isoleucine complex, in all the complex structures the loop connecting β3 and β4 moves and closes engulfing the ligand (Figure 5). This conformational change of loop β3β4 brings the residues Tyr167, Asp169 and Thr170 in proximity to the ligand and the membrane-distal subdomain adopts a closed conformation. The transition from an open to a closed state in the membrane-distal subdomain observed in all the Tlp3-LBD/ complex structures, disrupts two hydrogen bonds between the residue Tyr175, located in the moving loop, and the residues Asp217 and Ile222 from the loop connecting β6 and β7 at the membrane-proximal subdomain. As a result, the proximal subdomain adopts a slightly open conformation, inducing a downward displacement (~4 Å) of its C-terminal helix (Figure 5).

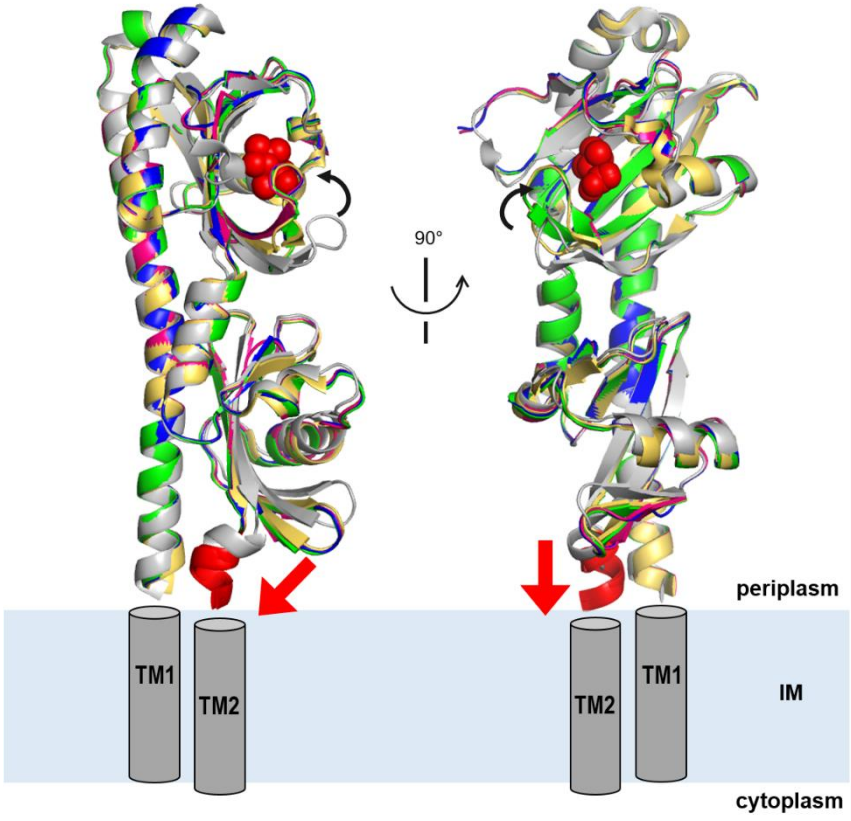


Figure 5. Conformational changes in the Tlp3 dCache LBD upon ligand binding. The superimposition of the subunit A of the free protein (shown in grey) and subunit B of the Tlp3-LBD/valine (shown in pink), Tlp3-LBD/leucine (shown in blue), Tlp3-LBD/α-amino-N-valeric acid (shown in green) and Tlp3-LBD/isoleucine (shown in yellow) shows the two extreme conformational states adopted by the Tlp3-LBD. Binding of the ligand as observed in all Tlp3-LBD/complex structures induces a downward ~4 Å displacement of the C-terminal helix (coloured red). Leucine, valine, isoleucine and α-amino-N-valeric acid are shown as sphere coloured red. IM: inner membrane.

Discussion

Previously, Tlp3 has been suggested as the *Campylobacter* chemoreceptor for multiple ligands (CcmL) that directly senses a broad range of signal molecules, including several amino acids: arginine, aspartate, glucosamine, isoleucine and lysine; tricarboxylic acid (TCA) cycle intermediates: malic acid, succinic acid and α-ketoglutarate, as well as purine and thiamine [1]. In the current study, these molecules were included in the thermal shift and ITC measurements to determine whether a direct interaction with Tlp3-LBD occurs. Interestingly, no binding between Tlp3-LBD and these molecules could be detected by thermal shift and ITC assays (Figure 2D, Table 4). Furthermore, analysis of high-resolution Fourier maps of Tlp3-LBD structures co-crystallised with L-lysine, L-arginine, L-aspartate or malic acid revealed no electron density that could be interpreted as any of these molecules bound to this dCache module. Together, these experimental data confirmed that except for isoleucine, the molecules previously reported as Tlp3 ligands do not directly bind to the dCache LBD of this chemoreceptor.

In the study performed by Rahman *et al.* [1], the K_D between Tlp3-LBD and the molecules mistakenly identified to directly bind this receptor was calculated by surface plasmon resonance (SPR). For these experiments, His_6-tagged Tlp3-LBD was loaded on a Ni^{2+} NTA sensor chip, after which several amino acids and salts of organic acid were independently loaded using single-cycle kinetics [1]. Although SPR is a technique frequently used to measure protein-ligand interactions [70]; this approach is not suitable to measure affinity data for very small analytes (Mr <1000 gr mol^{-1}), as the SPR signal is directly related to the change in mass of material binding to the sensor surface [70,71]. The molecular weight of the molecules

identified by Rahman *et al.* as Tlp3 ligands is <300 gr mol^{-1}, and thus the signal detected must have been very close to the noise signal which resulted in high rate of false positive hits.

As Tlp3-LBD ligand specificity required further investigation, in the present study, Tlp3-LBD was screened against a library of metabolic substrates that has been previously used in thermal shift assays to identify the specificity of chemoreceptors [45,49,54]. From this screening, seven compounds were found to induce significant changes in the Tm and were considered as hits. ITC measurements with the purified LBD of Tlp3 demonstrated direct binding of this sensory module with the amino acids: leucine, valine and α-amino-N-valeric acid. These findings are consistent with the structural and bioinformatic analysis discussed in Chapter 2, which established Tlp3 as a chemoreceptor that directly recognises amino acids [32].

The crystal structures of Tlp3-LBD in complex with the ligands identified here showed that they are bound at the membrane-distal subdomain. The ligand-binding pocket at membrane-distal subdomain has been involved in direct ligand recognition in almost all the previously characterised dCache sensor domains including Tlp3 [32] and other amino acid receptors; such as Mlp37 receptor from *V. parahaemolyticus* [47], McpB and McpC from *B. subtilis* [31,51], and Mlp24 from *V. cholerae* [50].

Detailed analysis of the different complexes revealed that the ligands bind to Tlp3-LBD in an almost identical mode to that shown in the Tlp3-LBD/isoleucine complex described in Chapter 2. After binding, the amino acid ligands are entirely shielded from the solvent, and the complexes are stabilised with the formation of numerous interactions between the ligands (both main- and side- chains) and the LBD. Similar to the Tlp3-LBD/isoleucine complex, the main chain of the leucine, valine and α-amino-N-valeric acid are interacting with five protein residues (Lys149, Trp151, Tyr167, Asp169 and Asp196) that are highly conserved across different amino acids receptors with a dCache fold [32]. In contrast, the aliphatic side chain of the amino acid ligands is perfectly fitted in a highly hydrophobic pocket (Tyr118, Val126, Leu128, Leu144 and Val171), which residues are not conserved in fact, their substitution have shown to determine the amino acid ligand specificity of the receptor [32,46,51]. Thus, our structural analysis showed that there is a correlation between the chemical nature of the ligand-binding pocket and the side chain of the amino acid ligands identified by Tlp3, and confirmed that Tlp3 is a receptor that recognises specifically branched-chain aliphatic amino acids (Leu, Val, Ile) through its membrane-distal subdomain.

Comparison of the Tlp3-LBD/ complexes showed no significative differences between the structures. However, two variable regions at the membrane-distal subdomain were identified, corresponding to the loops β2'α3 and β3β4 (Figure 6A). Structural superposition of the complex structures revealed that several residues from these variable loops move to accommodate the different side chain of the amino acid ligands recognised by Tlp3-LBD.

What is the physiological relevance of taxis towards branched amino acids in *C. jejuni*? *C. jejuni* is generally considered a non-saccharolytic bacterium, as most of the strains from this genus are unable to use glucose and other carbohydrates as substrates [39,72]. Therefore for optimal growth, *C. jejuni* is dependent on TCA and a limited number of amino acids (L-aspartate, L-glutamate, L-proline and L-serine) as primary carbon and energy sources [73,74]. Although *C. jejuni* cannot catabolise branched-chain amino acids, this pathogen has a dedicated import system for leucine, isoleucine and valine, known as the LIV (leucine, isoleucine and valine) branched-chain amino acid ABC transporter system. Six proteins comprise the LIV system: two periplasmic binding proteins LivJ and LivK, the permeases LivH and LivM, and the cytoplasmic ATPases LivG and LivF [75]. Furthermore, from these branched-chain amino acids *C. jejuni* cells have been shown to exhibit a chemoattractant response to isoleucine, which is mediated by Tlp3 receptor [1]. The uptake of branched-chain amino acids has been found to play a crucial role in *C. jejuni* intestinal colonisation since *liv* mutant strains had a severe colonisation defect of mouse and chicken animal models [17,75]. As branched-chain amino acids cannot be utilised as a carbon source by *C. jejuni*, their acquisition might be instead necessary for sustaining normal protein synthesis.

Conclusion

The structural and biophysical experiments performed here showed that the molecules previously reported as Tlp3 ligands (arginine, aspartate, glucosamine, lysine, malic acid, succinic acid, α-ketoglutarate, purine and thiamine) do not interact with the dCache LBD of this receptor. The results of this study do not support the notion that Tlp3 is a receptor for multiple ligands. On the contrary, direct ligand sensing by Tlp3 appears to be rather specific and limited to branched-chain aliphatic amino acids (isoleucine, leucine, valine and α-amino-N-valeric acid).

Supplementary information

Table S1. Tlp3-LBD thermal shift assay results with the compounds in Biolog PM1, PM3B and PM5 plates.

Table S1. Tlp3-LBD thermal shift assay results with the compounds from the Biolog PM1, PM3B and PM5 screens. The ligands that induced a stabilising effect are shown in bold.

PM1 plate (carbon sources)					
Well	Compound	ΔT_m (°C)	Well	Compound	ΔT_m (°C)
A1	Negative Control (water only)	0.0 ± 0.1	D3	D-Glucosaminic Acid *	-0.1 ± 0.4
A2	L-Arabinose	0.2 ± 1.1	D4	1,2-Propanediol	-0.5 ± 0.1
A3	N-Acetyl-D-Glucosamine	0.3 ± 0.5	D5	Tween 40	0.1 ± 0.1
A4	D-Saccharic Acid	0.0 ± 0.2	D6	α-Keto-Glutaric Acid*	-0.1 ± 0.0
A5	Succinic Acid *	-0.1 ± 0.6	D7	α-Keto-Butyric Acid	0.0 ± 0.6
A6	D-Galactose	0.0 ± 0.7	D8	α-Methyl-D-Galactoside	0.3 ± 0.4
A7	L-Aspartic Acid *	-0.2 ± 0.7	D9	α-D-Lactose	0.2 ± 0.5
A8	L-Proline	0.3 ± 0.3	D10	Lactulose	-0.2 ± 0.6
A9	D-Alanine	0.0 ± 0.2	D11	Sucrose	-0.4 ± 0.2
A10	D-Trehalose	0.2 ± 0.3	D12	Uridine	-0.5 ± 0.4
A11	D-Mannose	-0.1 ± 0.7	E1	L-Glutamine	0.3 ± 0.8
A12	Dulcitol	0.1 ± 0.1	E2	m-Tartaric Acid	-0.6 ± 0.3
B1	D-Serine	0.0 ± 0.5	E3	D-Glucose-1-Phosphate	-0.9 ± 0.2
B2	D-Sorbitol	0.2 ± 0.3	E4	D-Fructose-6-Phosphate	-0.8 ± 0.2
B3	Glycerol	-0.3 ± 0.7	E5	Tween 80	0.1 ± 0.1
B4	L-Fucose	0.1 ± 0.6	E6	α-Hydroxy Glutaric Acid-γ-Lactone	-0.7 ± 0.5
B5	D-Glucuronic Acid	-0.7 ± 0.5	E7	α-Hydroxybutyric Acid	0.2 ± 0.7
B6	D-Gluconic Acid	-0.7 ± 0.3	E8	β-Methyl-D-Glucoside	0.5 ± 0.7
B7	D,L-α-Glycerol-Phosphate	-1.1 ± 0.1	E9	Adonitol	-0.3 ± 0.3
B8	D-Xylose	-1.2 ± 0.3	E10	Maltotriose	-0.2 ± 0.3
B9	L-Lactic Acid	-0.7 ± 0.2	E11	2-Deoxy Adenosine	-0.3 ± 0.2
B10	Formic Acid	-0.5 ± 0.4	E12	Adenosine	-0.1 ± 0.5
B11	D-Mannitol	-0.7 ± 0.4	F1	Glycyl-L-Aspartic Acid	-0.6 ± 0.4
B12	L-Glutamic Acid	-0.7 ± 0.0	F2	Citric Acid	-0.3 ± 0.7
C1	D-Glucose-6-Phosphate	-0.9 ± 0.5	F3	m-Inositol	-0.3 ± 0.7
C2	D-Galactonic Acid-γ-Lactone	-1.1 ± 0.4	F4	D-Threonine	-0.6 ± 0.2
C3	D,L-Malic Acid *	-0.6 ± 0.6	F5	Fumaric Acid	-0.9 ± 0.6
C4	D-Ribose	-0.7 ± 0.5	F6	Bromo Succinic Acid	-1.4 ± 0.7
C5	Tween 20	-0.2 ± 0.1	F7	Propionic Acid	-0.9 ± 0.2
C6	L-Rhamnose	-0.8 ± 0.3	F8	Mucic Acid	-0.8 ± 0.4
C7	D-Fructose	-0.7 ± 0.7	F9	Glycolic Acid	-0.7 ± 0.9
C8	Acetic Acid	0.0 ± 0.3	F10	Glyoxylic Acid	-0.7 ± 0.1
C9	α-D-Glucose	-0.1 ± 0.3	F11	D-Cellobiose	-1.0 ± 0.5
C10	Maltose	0.2 ± 0.4	F12	Inosine	-1.2 ± 0.3
C11	D-Melibiose	-0.1 ± 0.4	G1	Glycyl-L-Glutamic Acid	-0.7 ± 1.0
C12	Thymidine	0.2 ± 0.2	G2	Tricarballylic Acid	-1.1 ± 1.0
D1	L-Asparagine	-0.2 ± 0.1	G3	L-Serine	-1.1 ± 0.2
D2	D-Aspartic Acid *	-0.2 ± 0.4	G4	L-Threonine	-0.7 ± 0.6

Continuation Table S1

PM1 plate (carbon sources)					
Well	Compound	ΔTm (°C)	Well	Compound	ΔTm (°C)
G5	L-Alanine	-0.6 ± 0.7	H3	m-Hydroxy Phenyl Acetic Acid	-1.3 ± 0.5
G6	L-Alanyl-Glycine	-0.6 ± 0.4	H4	Tyramine	-0.8 ± 0.7
G7	Acetoacetic Acid	-0.8 ± 0.1	H5	D-Psicose	-0.9 ± 0.4
G8	N-Acetyl-β-D-Mannosamine	-1.1 ± 0.2	H6	L-Lyxose	-1.3 ± 0.2
G9	Mono Methyl Succinate	-0.8 ± 0.6	H7	Glucuronamide	-0.5 ± 1.2
G10	Methyl Pyruvate	-1.8 ± 0.0	H8	Pyruvic Acid	0.1 ± 0.2
G11	D-Malic Acid *	-1.2 ± 0.8	H9	L-Galactonic Acid-γ-Lactone	0.4 ± 0.3
G12	L-Malic Acid *	-0.8 ± 0.4	H10	D-Galacturonic Acid	0.4 ± 0.4
H1	Glycyl-L-Proline	-1.2 ± 0.2	H11	Phenylethylamine	0.1 ± 0.1
H2	p-Hydroxy Phenyl Acetic Acid	-1.4 ± 0.8	H12	2-Aminoethanol	-0.3 ± 0.1
PM3 plate (nitrogen sources)					
A1	Negative control (water only)	0.0 ± 0.1	C5	D-Aspartic Acid *	-0.7 ± 0.2
A2	Ammonia	-0.4 ± 0.9	C6	D-Glutamic Acid	-0.9 ± 0.2
A3	Nitrite	-0.2 ± 0.2	C7	D-Lysine *	-0.4 ± 0.4
A4	Nitrate	-1.0 ± 0.3	C8	D-Serine	-0.1 ± 0.4
A5	Urea	-0.8 ± 0.1	C9	D-Valine	0.1 ± 0.3
A6	Biuret	-0.9 ± 0.2	C10	L-Citrulline	0.2 ± 0.1
A7	L-Alanine	-0.8 ± 0.3	C11	L-Homoserine	-0.2 ± 0.4
A8	L-Arginine *	-1.0 ± 0.3	C12	L-Ornithine	-0.1 ± 0.2
A9	L-Asparagine	-1.1 ± 0.1	D1	N-Acetyl-L-Glutamic Acid	0.9 ± 0.4
A10	L-Aspartic Acid *	-0.4 ± 0.3	D2	N-Phthaloyl-L-Glutamic Acid	-0.3 ± 0.2
A11	L-Cysteine	-0.5 ± 0.9	D3	L-Pyroglutamic Acid	0.0 ± 0.7
A12	L-Glutamic Acid	0.2 ± 0.5	D4	Hydroxylamine	-0.4 ± 0.2
B1	L-Glutamine	-0.8 ± 0.5	D5	Methylamine	-0.5 ± 0.3
B2	Glycine	-0.4 ± 0.3	D6	N-Amylamine	0.1 ± 0.4
B3	L-Histidine	-0.4 ± 0.2	D7	N-Butylamine	0.0 ± 0.5
B4	**L- Isoleucine**	**2.1 ± 0.8**	D8	Ethylamine	0.3 ± 0.4
B5	**L-Leucine**	**2.4 ± 0.2**	D9	Ethanolamine	0.0 ± 0.3
B6	L-Lysine *	-0.7 ± 0.0	D10	Ethylenediamine	-0.2 ± 0.2
B7	L-Methionine	-0.3 ± 0.3	D11	Putrescine	-0.5 ± 0.5
B8	L-Phenylalanine	-0.2 ± 0.0	D12	Agmatine	-0.8 ± 0.8
B9	L-Proline	-0.9 ± 0.2	E1	Histamine	-0.9 ± 0.6
B10	L-Serine	-0.9 ± 0.3	E2	β-Phenylethylamine	-1.2 ± 0.2
B11	L-Threonine	-0.8 ± 0.2	E3	Tyramine	-0.4 ± 0.3
B12	L-Tryptophan	0.3 ± 0.5	E4	Acetamide	-0.1 ± 0.4
C1	L-Tyrosine	0.0 ± 0.5	E5	Formamide	-0.4 ± 0.5
C2	**L-Valine**	**1.3 ± 0.1**	E6	Glucuronamide	-0.4 ± 0.5
C3	D-Alanine	-0.5 ± 0.7	E7	D,L-Lactamide	-0.1 ± 0.8
C4	D-Asparagine	-0.2 ± 0.6	E8	D-Glucosamine*	-0.4 ± 0.3

Continuation Table S1

PM3 plate (nitrogen sources)					
Well	Compound	ΔT_m (°C)	Well	Compound	ΔTm (°C)
E9	D-Galactosamine	-1.4 ± 1.2	G5	Allantoin	-0.1 ± 0.5
E10	D-Mannosamine	-0.3 ± 0.3	G6	Parabanic Acid	-0.3 ± 0.3
E11	N-Acetyl-D-Glucosamine	0.3 ± 0.3	G7	D,L-α-Amino-N-Butyric Acid	0.0 ± 0.1
E12	N-Acetyl-D-Galactosamine	-0.6 ± 0.1	G8	γ-Amino-N-Butyric Acid	0.1 ± 0.2
F1	N-Acetyl-D-Mannosamine	0.1 ± 0.5	G9	ε-Amino -N-Caproic Acid	0.1 ± 0.2
F2	Adenine	-0.4 ± 0.4	G10	D,L-α-Amino-Caprylic Acid	0.2 ± 0.8
F3	Adenosine	-0.8 ± 0.2	G11	δ-Amino-N-Valeric Acid	-0.4 ± 0.2
F4	Cytidine	-0.1 ± 1.2	G12	**α-Amino-N-Valeric Acid**	**1.4 ± 0.3**
F5	Cytosine	-0.3 ± 0.7	H1	Ala -Asp	-0.4 ± 0.3
F6	Guanine	-1.1 ± 0.2	H2	Ala -Gln	-0.1 ± 0.7
F7	Guanosine	-0.3 ± 0.4	H3	Ala -Glu	0.2 ± 0.3
F8	Thymine	0.2 ± 0.9	H4	Ala -Gly	-0.6 ± 0.2
F9	Thymidine	-0.2 ± 0.3	H5	Ala -His	-0.3 ± 0.2
F10	Uracil	-1.1 ± 0.2	H6	Ala -Leu	0.1 ± 0.2
F11	Uridine	-0.4 ± 0.4	H7	**Ala -Thr**	**1.1 ± 0.3**
F12	Inosine	-0.1 ± 0.9	H8	Gly -Asn	-0.3 ± 0.4
G1	Xanthine	-0.5 ± 0.2	H9	Gly -Gln	-0.4 ± 0.4
G2	Xanthosine	-0.1 ± 0.5	H10	Gly -Glu	-0.5 ± 0.1
G3	Uric Acid	-0.2 ± 0.1	H11	Gly -Met	-0.3 ± 0.0
G4	Alloxan	0.0 ± 0.9	H12	Met -Ala	0.0 ± 0.3
PM5 plate (nutrient supplements)					
A1	Negative control (water only)	0.0 ± 0.1	B6	L-Lysine *	-0.5 ± 0.1
A3	L-Alanine	-0.5 ± 0.1	B7	L-Methionine	0.3 ± 0.5
A4	L-Arginine *	0.1 ± 0.2	B8	L-Phenylalanine	-0.3 ± 0.4
A5	L-Asparagine	-0.2 ± 0.6	B9	Guanosine-3',5'-cyclic monophosphate	0.2 ± 0.5
A6	L-Aspartic Acid *	-0.7 ± 0.3	B10	Guanine	-0.3 ± 0.6
A7	L-Cysteine	-0.2 ± 0.4	B11	Guanosine	-0.4 ± 0.1
A8	L-Glutamic Acid	0.0 ± 0.8	B12	2'-Deoxy Guanosine	0.0 ± 0.3
A9	Adenosine-3',5'-cyclicmonophosphate	0.0 ± 0.5	C1	L-Proline	-0.6 ± 0.1
A10	Adenine	-0.3 ± 0.6	C2	L-Serine	-0.3 ± 0.3
A11	Adenosine	0.1 ± 0.6	C3	L-Threonine	-0.4 ± 0.4
A12	2'-Deoxy Adenosine	-0.2 ± 0.1	C4	L-Tryptophan	0.0 ± 0.2
B1	L-Glutamine	-0.2 ± 0.3	C5	L-Tyrosine	-0.2 ± 0.5
B2	Glycine	-0.3 ± 0.3	C6	L-Valine	0.7 ± 0.4
B3	L-Histidine	0.2 ± 0.5	C7	L-isoleucine + L-valine	-0.3 ± 0.2
B4	L- Isoleucine	1.1 ± 0.1	C8	trans-4-Hydroxy L-Proline	-0.7 ± 0.3
B5	L-Leucine	0.6 ± 0.3	C9	(5) 4-Amino-Imidazole-4(5)-Carboxamide	-0.6 ± 0.6

Continuation Table S1

PM5 plate (nutrient supplements)					
Well	Compound	ΔT_m (°C)	Well	Compound	ΔT_m (°C)
C10	Hypoxanthine	-0.5 ± 0.3	F6	Hematin	-0.2 ± 0.2
C11	Inosine	-0.3 ± 0.3	F7	Deferoxamine Mesylate	-0.3 ± 0.2
C12	2'-Deoxy Inosine	-1.0 ± 0.3	F8	D-(+)-Glucose	-0.2 ± 0.1
D1	L-Ornithine	-0.7 ± 0.5	F9	N-Acetyl D-Glucosamine	-0.4 ± 0.2
D2	L-Citrulline	-0.4 ± 0.2	F10	Thymine	-0.3 ± 0.1
D3	Chorismic Acid	-1.0 ± 0.1	F11	Glutathione (reduced form)	0.1 ± 0.4
D4	(-)Shikimic Acid	-1.1 ± 0.2	F12	Thymidine	-0.1 ± 0.1
D5	L-Homoserine Lactone	-0.6 ± 0.1	G1	Oxaloacetic Acid	-0.3 ± 0.2
D6	D-Alanine	-0.4 ± 0.1	G2	D-Biotin	-0.3 ± 0.1
D7	D-Aspartic Acid *	-0.5 ± 0.1	G3	Cyano-Cobalamine	-0.3 ± 0.2
D8	D-Glutamic Acid	-0.8 ± 0.2	G4	p-Amino-Benzoic Acid	-0.1 ± 0.2
D9	D,L-α,ε-Diaminopimelic Acid	-0.7 ± 0.4	G5	Folic Acid	-0.7 ± 0.4
D10	Cytosine	-0.7 ± 0.1	G6	Inosine + Thiamine	-0.7 ± 0.3
D11	Cytidine	-0.4 ± 0.4	G7	Thiamine *	-0.1 ± 0.4
D12	**2'-Deoxy Cytidine**	**1.4 ± 0.8**	G8	Thiamine Pyrophosphate	0.0 ± 0.0
E1	Putrescine	-0.3 ± 0.1	G9	Riboflavin	-0.1 ± 0.2
E2	Spermidine	-0.2 ± 0.0	G10	Pyrrolo-Quinoline Quinone	-0.1 ± 0.5
E3	Spermine	-0.4 ± 0.3	G11	Menadione	-0.5 ± 0.7
E4	Pyridoxine	0.0 ± 0.2	G12	m-Inositol	-0.2 ± 0.4
E5	Pyridoxal	0.0 ± 0.4	H1	Butyric Acid	-0.4 ± 0.2
E6	Pyridoxamine	-0.4 ± 0.3	H2	D,L-α-Hydroxy-Butyric Acid	-0.4 ± 0.3
E7	β-Alanine	-0.4 ± 0.2	H3	α-Keto-Butyric Acid	0.0 ± 0.3
E8	D-Pantothenic Acid	-0.4 ± 0.3	H4	Caprylic Acid	-0.4 ± 0.8
E9	Orotic Acid	0.1 ± 0.3	H5	D,L-α-Lipoic Acid (oxidized form)	-0.2 ± 0.4
E10	Uracil	-0.3 ± 0.3	H6	D,L-Mevalonic Acid	0.0 ± 0.6
E11	Uridine	0.0 ± 0.6	H7	D,L-Carnitine	-0.1 ± 0.6
E12	2'-Deoxy Uridine	-0.1 ± 0.2	H8	Choline	-0.1 ± 0.3
F1	Quinolinic Acid	0.1 ± 0.3	H9	Tween 20	-0.2 ± 0.1
F2	Nicotinic Acid	0.1 ± 0.2	H10	Tween 40	0.1 ± 0.3
F3	Nicotinamide	0.2 ± 0.2	H11	Tween 60	-0.1 ± 0.1
F4	β-Nicotilamide Adenine Dinucleotide	-0.4 ± 0.1	H12	Tween 80	0.1 ± 0.2
F5	δ-Amino-Levulinic Acid	-0.1 ± 0.2			

* Chemical molecules previously suggested as Tlp3 ligands [1]

References

1 Rahman, H. *et al.* Characterisation of a multi-ligand binding chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. *PLoS Pathog* **10**, e1003822, doi:10.1371/journal.ppat.1003822 (2014).

2 Li, Z. *et al.* Methyl-accepting chemotaxis proteins 3 and 4 are responsible for *Campylobacter jejuni* chemotaxis and jejuna colonization in mice in response to sodium deoxycholate. *J Med Microbiol* **63**, 343-354, doi:10.1099/jmm.0.068023-0 (2014).

3 Boes, J. *et al.* Prevalence and diversity of Campylobacter jejuni in pig herds on farms with and without cattle or poultry. *J Food Prot* **68**, 722-727 (2005).

4 Oporto, B., Esteban, J. I., Aduriz, G., Juste, R. A. & Hurtado, A. Prevalence and strain diversity of thermophilic campylobacters in cattle, sheep and swine farms. *J Appl Microbiol* **103**, 977-984, doi:10.1111/j.1365-2672.2007.03328.x (2007).

5 Beery, J. T., Hugdahl, M. B. & Doyle, M. P. Colonization of gastrointestinal tracts of chicks by Campylobacter jejuni. *Appl Environ Microbiol* **54**, 2365-2370 (1988).

6 Petersen, L., Nielsen, E. M., Engberg, J., On, S. L. & Dietz, H. H. Comparison of genotypes and serotypes of Campylobacter jejuni isolated from Danish wild mammals and birds and from broiler flocks and humans. *Appl Environ Microbiol* **67**, 3115-3121, doi:10.1128/AEM.67.7.3115-3121.2001 (2001).

7 Shane, S. M. Campylobacter infection of commercial poultry. *Rev Sci Tech* **19**, 376-395 (2000).

8 Dingle, K. E. *et al.* Molecular characterization of Campylobacter jejuni clones: a basis for epidemiologic investigation. *Emerg Infect Dis* **8**, 949-955, doi:10.3201/eid0809.020122 (2002).

9 Friedman, C. R. *et al.* Risk factors for sporadic Campylobacter infection in the United States: A case-control study in FoodNet sites. *Clin Infect Dis* **38 Suppl 3**, S285-296, doi:10.1086/381598 (2004).

10 Epps, S. V. *et al.* Foodborne Campylobacter: infections, metabolism, pathogenesis and reservoirs. *Int J Environ Res Public Health* **10**, 6292-6304, doi:10.3390/ijerph10126292 (2013).

11 Crim, S. M. *et al.* Incidence and trends of infection with pathogens transmitted commonly through food--Foodborne Diseases Active Surveillance Network, 10 U.S. sites, 2006-2013. *MMWR Morb Mortal Wkly Rep* **63**, 328-332 (2014).

12 Blaser, M. J. Epidemiologic and clinical features of Campylobacter jejuni infections. *J Infect Dis* **176 Suppl 2**, S103-105 (1997).

13 Friedman, C., Neimann, J., Wegener, H. & Tauxe, R. in *Campylobacter* Vol. 2nd (eds I Nachamkin & Blaser M) 121-138 (ASM Press, 2000).

14 Schmidt-Ott, R. *et al.* Improved serological diagnosis stresses the major role of Campylobacter jejuni in triggering Guillain-Barre syndrome. *Clin Vaccine Immunol* **13**, 779-783, doi:10.1128/CVI.00065-06 (2006).

15 Dasti, J. I., Tareen, A. M., Lugert, R., Zautner, A. E. & Gross, U. *Campylobacter jejuni*: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. *Int J Med Microbiol* **300**, 205-211, doi:10.1016/j.ijmm.2009.07.002 (2010).

16 Josenhans, C. & Suerbaum, S. The role of motility as a virulence factor in bacteria. *Int J Med Microbiol* **291**, 605-614, doi:10.1078/1438-4221-00173 (2002).

17 Hendrixson, D. R. & DiRita, V. J. Identification of Campylobacter jejuni genes involved in commensal colonization of the chick gastrointestinal tract. *Mol Microbiol* **52**, 471-484, doi:10.1111/j.1365-2958.2004.03988.x (2004).

18 Morooka, T., Umeda, A. & Amako, K. Motility as an intestinal colonization factor for *Campylobacter jejuni*. *J Gen Microbiol* **131**, 1973-1980, doi:10.1099/00221287-131-8-1973 (1985).

19 Tareen, A. M., Dasti, J. I., Zautner, A. E., Gross, U. & Lugert, R. *Campylobacter jejuni* proteins Cj0952c and Cj0951c affect chemotactic behaviour towards formic acid and are important for invasion of host cells. *Microbiology* **156**, 3123-3135, doi:10.1099/mic.0.039438-0 (2010).

20 Zautner, A. E., Tareen, A. M., Gross, U. & Lugert, R. Chemotaxis in *Campylobacter jejuni*. *European journal of microbiology & immunology* **2**, 24-31, doi:10.1556/EuJMI.2.2012.1.5 (2012).

21 Zhulin, I. B. The superfamily of chemotaxis transducers: from physiology to genomics and back. *Adv Microb Physiol* **45**, 157-198 (2001).

22 Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends in biochemical sciences* **33**, 9-19, doi:10.1016/j.tibs.2007.09.014 (2008).

23 Wuichet, K., Alexander, R. P. & Zhulin, I. B. Comparative genomic and protein sequence analyses of a complex system controlling bacterial chemotaxis. *Methods in enzymology* **422**, 1-31, doi:10.1016/S0076-6879(06)22001-9 (2007).

24 Lacal, J., Garcia-Fontana, C., Munoz-Martinez, F., Ramos, J. L. & Krell, T. Sensing of environmental signals: classification of chemoreceptors according to the size of their ligand binding regions. *Environ. Microbiol.* **12**, 2873-2884, doi:10.1111/j.1462-2920.2010.02325.x (2010).

25 Milburn, M. V. *et al.* Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* **254**, 1342-1347 (1991).

26 Kim, K. K., Yokota, H. & Kim, S. H. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* **400**, 787-792, doi:10.1038/23512 (1999).

27 Goers Sweeney, E. *et al.* Structure and proposed mechanism for the pH-sensing *Helicobacter pylori* chemoreceptor TlpB. *Structure* **20**, 1177-1188, doi:10.1016/j.str.2012.04.021 (2012).

28 Upadhyay, A. A., Fleetwood, A. D., Adebali, O., Finn, R. D. & Zhulin, I. B. Cache Domains That are Homologous to, but Different from PAS Domains Comprise the Largest Superfamily of Extracellular Sensors in Prokaryotes. *PLoS Comput. Biol.* **12**, e1004862; 1004810.1001371/journal.pcbi.1004862, doi:10.1371/journal.pcbi.1004862 (2016).

29 Pineda-Molina, E. *et al.* Evidence for chemoreceptors with bimodular ligand-binding regions harboring two signal-binding sites. *Proc. Natl. Acad. Sci. U S A* **109**, 18926-18931, doi:10.1073/pnas.1201400109 (2012).

30 Martin-Mora, D. *et al.* McpQ is a specific citrate chemoreceptor that responds preferentially to citrate/metal ion complexes. *Environ. Microbiol.* **18**, 3284-3295, doi:10.1111/1462-2920.13030 (2016).

31 Glekas, G. D. *et al.* The *Bacillus subtilis* chemoreceptor McpC senses multiple ligands using two discrete mechanisms. *J. Biol. Chem.* **287**, 39412-39418, doi:10.1074/jbc.M112.413518 (2012).

32 Liu, Y. C., Machuca, M. A., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. *Acta crystallographica. Section D, Biological crystallography* **71**, 2127-2136, doi:10.1107/S139900471501384X (2015).

33 Machuca, M. A., Liu, Y. C., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni* Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition. *J. Struct. Biol.* **194**, 205-213, doi:10.1016/j.jsb.2016.02.019 (2016).

34 Matilla, M. A. & Krell, T. Chemoreceptor-based signal sensing. *Curr. Opin. Biotechnol.* **45**, 8-14, doi:10.1016/j.copbio.2016.11.021 (2017).

35 Machuca, M. A. *et al.* Helicobacter pylori chemoreceptor TlpC mediates chemotaxis to lactate. *Sci Rep* **7**, 14089, doi:10.1038/s41598-017-14372-2 (2017).

36 Mund, N. L. *et al.* Association of Campylobacter Jejuni ssp. Jejuni Chemotaxis Receptor Genes with Multilocus Sequence Types and Source of Isolation. *European journal of microbiology & immunology* **6**, 162-177, doi:10.1556/1886.2015.00041 (2016).

37 Marchant, J., Wren, B. & Ketley, J. Exploiting genome sequence: predictions for mechanisms of Campylobacter chemotaxis. *Trends Microbiol* **10**, 155-159 (2002).

38 Day, C. J. *et al.* Variation of chemosensory receptor content of Campylobacter jejuni strains and modulation of receptor gene expression under different in vivo and in vitro growth conditions. *BMC Microbiol* **12**, 128, doi:10.1186/1471-2180-12-128 (2012).

39 Parkhill, J. *et al.* The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665-668, doi:10.1038/35001088 (2000).

40 Hugdahl, M. B., Beery, J. T. & Doyle, M. P. Chemotactic behavior of *Campylobacter jejuni*. *Infection and immunity* **56**, 1560-1566 (1988).

41 Day, C. J. *et al.* A direct-sensing galactose chemoreceptor recently evolved in invasive strains of Campylobacter jejuni. *Nat Commun* **7**, 13206, doi:10.1038/ncomms13206 (2016).

42 Vegge, C. S., Brondsted, L., Li, Y. P., Bang, D. D. & Ingmer, H. Energy taxis drives Campylobacter jejuni toward the most favorable conditions for growth. *Appl Environ Microbiol* **75**, 5308-5314, doi:10.1128/AEM.00287-09 (2009).

43 Hartley-Tassell, L. E. *et al.* Identification and characterization of the aspartate chemosensory receptor of *Campylobacter jejuni*. *Mol Microbiol* **75**, 710-730, doi:10.1111/j.1365-2958.2009.07010.x (2010).

44 Day, C. *et al.* Characterisation of the galactose chemosensory receptor of *Campylobacter jejuni*. *In: Proceedings of The Australian Society for Microbiology Annual Scientific Meeting. Australian Society for Microbiology.* **Abstract #107.** (2014).

45 McKellar, J. L., Minnell, J. J. & Gerth, M. L. A high-throughput screen for ligand binding reveals the specificities of three amino acid chemoreceptors from *Pseudomonas syringae* pv. *actinidiae*. *Mol. Microbiol.* **96**, 694-707, doi:10.1111/mmi.12964 (2015).

46 Webb, B. A., Hildreth, S., Helm, R. F. & Scharf, B. E. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Appl. Environ. Microbiol.* **80**, 3404-3415, doi:10.1128/AEM.00115-14 (2014).

47 Nishiyama, S. *et al.* Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci. Rep.* **6**, 20866; 20810.21038/srep20866, doi:10.1038/srep20866 (2016).

48 Liu, X., Wood, P. L., Parales, J. V. & Parales, R. E. Chemotaxis to pyrimidines and identification of a cytosine chemoreceptor in *Pseudomonas putida*. *Journal of bacteriology* **191**, 2909-2916, doi:10.1128/JB.01708-08 (2009).

49 Fernandez, M., Morel, B., Corral-Lugo, A. & Krell, T. Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol. Microbiol.* **99**, 34-42, doi:10.1111/mmi.13215 (2016).

50 Nishiyama, S. *et al.* Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infection and immunity* **80**, 3170-3178, doi:10.1128/IAI.00039-12 (2012).

51 Glekas, G. D. *et al.* A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. *J. Biol. Chem.* **285**, 1870-1878, doi:10.1074/jbc.M109.072108 (2010).

52 Dwivedi, R. *et al.* L-fucose influences chemotaxis and biofilm formation in Campylobacter jejuni. *Mol Microbiol* **101**, 575-589, doi:10.1111/mmi.13409 (2016).

53 Fernandez, M., Morel, B., Corral-Lugo, A. & Krell, T. Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol Microbiol* **99**, 34-42, doi:10.1111/mmi.13215 (2016).

54 Corral-Lugo, A. *et al.* Assessment of the contribution of chemoreceptor-based signalling to biofilm formation. *Environ Microbiol* **18**, 3355-3372, doi:10.1111/1462-2920.13170 (2016).

55 Machuca, M. A., Liu, Y. C., Beckham, S. A. & Roujeinikova, A. Cloning, refolding, purification and preliminary crystallographic analysis of the sensory domain of the *Campylobacter* chemoreceptor for multiple ligands (CcmL). *Acta crystallographica. Section F, Structural biology communications* **71**, 211-216, doi:10.1107/S2053230X1500045X (2015).

56 Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221, doi:10.1038/nprot.2007.321 (2007).

57 Lo, M. C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* **332**, 153-159, doi:10.1016/j.ab.2004.04.031 (2004).

58 Pantoliano, M. W. *et al.* High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* **6**, 429-440, doi:10.1177/108705710100600609 (2001).

59 Bochner, B. R., Gadzinski, P. & Panomitros, E. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Res* **11**, 1246-1255, doi:10.1101/gr.186501 (2001).

60 McPhillips, T. M. *et al.* Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* **9**, 401-406 (2002).

61 Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography* **67**, 271-281, doi:10.1107/S0907444910048675 (2011).

62 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography* **69**, 1204-1214, doi:10.1107/S0907444913000061 (2013).

63 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).

64 McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**, 458-464, doi:10.1107/S0907444905001617 (2005).

65 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

66 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

67 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).

68 Pence, H. E. & Williams, A. ChemSpider: An Online Chemical Information Resource. *J. Chem. Educ.* **87** 1123-1124, doi:10.1021/ed100697w (2010).

69 Dundas, J. *et al.* CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, W116-118, doi:10.1093/nar/gkl282 (2006).

70 Van Der Merwe, P. A. Surface plasmon resonance. *Protein-ligand interactions: hydrodynamics and calorimetry* **1**, 137-170 (2001).

71 Liedberg, B., Nylander, C. & Lundstrom, I. Biosensing with surface plasmon resonance--how it all started. *Biosens Bioelectron* **10**, i-ix (1995).

72 Muraoka, W. T. & Zhang, Q. Phenotypic and genotypic evidence for L-fucose utilization by Campylobacter jejuni. *J Bacteriol* **193**, 1065-1075, doi:10.1128/JB.01252-10 (2011).

73 Kelly, D. J. The physiology and metabolism of Campylobacter jejuni and Helicobacter pylori. *Symposium series*, 16S-24S (2001).

74 Stahl, M., Butcher, J. & Stintzi, A. Nutrient acquisition and metabolism by Campylobacter jejuni. *Front Cell Infect Microbiol* **2**, 5, doi:10.3389/fcimb.2012.00005 (2012).

75 Ribardo, D. A. & Hendrixson, D. R. Analysis of the LIV system of Campylobacter jejuni reveals alternative roles for LivJ and LivK in commensalism beyond branched-chain amino acid transport. *J Bacteriol* **193**, 6233-6243, doi:10.1128/JB.05473-11 (2011).

Chapter 6: General discussion and future directions

Summary of the main findings and conclusions

In *Campylobacter jejuni* and *Helicobacter pylori*, chemotaxis is an essential colonisation factor that allows them to reach their preferred infection site within their hosts [1-3]. During this process, environmental stimuli are detected by the extracytoplasmic ligand binding domains (LBDs) of membrane-embedded transducer-like proteins (Tlps). LBDs of chemoreceptors are largely diverse in amino acid sequence and three-dimensional structure [4,5] and have been shown to employ different mechanisms for ligand recognition [6].

In the present work the LBD of three chemoreceptors, Tlp1 and Tlp3 from *C. jejuni*, and TlpC from *H. pylori* have been characterised using a combination of X-ray crystallography, biochemistry and biophysical tools. The crystal structure of these LBDs revealed that all three adopt a double Cache (dCache) fold with two tandem PAS-like subdomains (membrane-proximal and membrane-distal) folding against a long stalk helix. dCache domains are part of the recently described Cache superfamily, which is the most abundant domain among extracellular sensory receptors in prokaryotes [4,5].

The detailed analysis of the crystal structure of dCache LBD of Tlp3 (Tlp3-LBD) and its complex with an attractant ligand (isoleucine) is described in **Chapter 2**. The crystal structure of Tlp3-LBD in complex with isoleucine (Tlp3-LBD/isoleucine) revealed that the ligand is recognised through the membrane-distal subdomain of this dCache module. In most of the previously characterised dCache LBDs, direct sensing involves binding of their cognate ligands to the distal subdomain [7-10]. Tlp3-LBD/isoleucine complex is stabilised by extensive interactions between several Tlp3-LBD residues with both the side and main chain moieties of isoleucine.

This elucidated crystallographic structure is further evinced by isothermal titration calorimetry (ITC) experiments on Tlp3-LBD variants. Mutagenesis experiments at the membrane-distal subdomain showed that any substitution of the residues that are stabilising the amino and carboxyl groups of the ligand (Lys149, Trp151, Tyr167, Asp169 and Asp196) had a negative effect on isoleucine binding. Interestingly, these residues are strongly conserved in other dCache sensory modules from different bacterial species and constitute a consensus motif (**D***XXX*(**R**/**K**)*X***WY***XX***A**), which is suggested here as a signature for amino acid receptors in a wide range of bacteria and archaea species.

Previously, Tlp3 has been suggested as the *Campylobacter* chemoreceptor for multiple ligands (CcmL) [11] that directly senses a broad range of signal molecules, including several amino acids, tricarboxylic acid (TCA) cycle intermediates, nucleotides and thiamine [11]. In **Chapter 5**, several biophysical and structural data demonstrated that with the exception of isoleucine, the molecules previously reported as Tlp3 ligands do not directly bind to the dCache LBD of this chemoreceptor. On the contrary, direct ligand sensing by Tlp3 appears to be rather specific and limited to branched-chain aliphatic amino acids.

Besides isoleucine, Tlp3-LBD was found to directly bind the amino acids leucine, valine and α-amino-N-valeric acid. The crystal structure of Tlp3-LBD in complex with these amino acids revealed they are bound at the distal subdomain in an almost identical mode to that in the Tlp3-LBD/isoleucine complex. Comparison of the membrane-distal subdomain of the four ligand-bound structures identified two flexible regions corresponding to the β2'α3 and β3β4 connecting loops. This structural flexibility of the distal subdomain enables Tlp3-LBD to accommodate the different branched-chain aliphatic amino acid ligands at the same binding site.

The present study also provides evidence for a possible signal transduction mechanism which might be common in chemoreceptors from the dCache superfamily. Conformational changes induced by ligand binding to the receptor suggests that dCache Tlp3 transmits the signal across the membrane into the cytoplasm by a piston displacement mechanism. This piston-like movement has been previously described as one of the signalling mechanisms for the Tar receptor from *E. coli* and *Salmonella typhimurium* [12,13]. Tar receptor's LBD has a four-helix bundle (4HB) fold [14,15], which has notable structural differences with the dCache domains [8,10]. Despite these structural differences, **Chapter 2** and **Chapter 5** show preliminary evidence that both chemoreceptor families might use a similar signal transduction process.

Chapter 3 describes the detailed biophysical and structural analyses of the LBD of *C. jejuni* receptor Tlp1 (Tlp1-LBD), which mediates taxis towards aspartic acid [16]. Despite the implication of Tlp1 sensing this amino acid, high-resolution Fourier maps derived from Tlp1-crystals grown in the presence of L-aspartate, did not present an electron density that could be interpreted as this amino acid bound to Tlp1-LBD. Furthermore, inspection of Tlp1 membrane-distal subdomain, which has been implicated in the binding of small-molecule ligands in Tlp3 (Chapter 2 and Chapter 5) and other dCache-containing proteins [9,10,17], revealed the lack of a ligand-binding pocket. Concordantly, ITC experiments showed no

binding between Tlp1-LBD and aspartate. These results demonstrated that the mechanism of aspartate recognition by Tlp1-LBD does not involve direct ligand-receptor interaction and suggest that Tlp1-LBD is likely to recognise this amino acid indirectly, for instance *via* an as-yet-unidentified periplasmic binding protein (PBP).

Detection of the chemoeffector signal by binding to chemoeffector-engaged PBPs have been observed for several receptors [18-25] including the dCache McpC chemoreceptor from *Bacillus subtilis* [9]. McpC receptor senses several amino acids indirectly by the recognition of four binding proteins, which interact with its membrane-proximal subdomain [9]. In contrast, Tlp1-LBD structure suggests that the PBP docking site is located on the membrane-distal subdomain, instead of the proximal. This is consistent with the analysis of the distribution of conservative substitutions in Tlp1-LBD across different *C. jejuni* strains showing that most strain-to-strain variations are located at the membrane-proximal subdomain and that the surface of the membrane-distal subdomain is highly conserved (**Chapter 3**) [26].

As a first step towards elucidating TlpC role in chemotaxis, in **Chapter 4** the crystal structure of this receptor's LBD (TlpC-LBD) was determined and analysed. Interestingly, analysis of TlpC-LBD structure revealed the presence of a non-protein molecule bound to the receptor. The shape of the electron density observed did not match any of the components used in the purification or crystallisation buffers. Therefore, it was hypothesised that the ligand trapped in the crystal could be a product of proteolytic degradation of the TlpC or a metabolite produced during cell growth. The co-purified molecule was identified as lactate, which is the first ligand reported to be recognised by TlpC receptor [27].

The chemotactic behaviour of wild-type *H. pylori* cells was evaluated in the presence of lactate. Experiments showed that this organic compound triggers a typical chemoattractant response, which was drastically reduced in an isogenic *ΔtlpC* mutant. These results demonstrated that *H. pylori* seeks out lactate by using chemotaxis and that TlpC is the primary chemoreceptor for lactate sensing. Indeed, lactate can be utilised by *H. pylori* as a carbon or energy source [28] and has been previously shown to promote its growth *in vitro* and *in vivo* [29].

Analysis of TlpC-LBD/lactate complex structure showed that lactate is bound at the putative binding site of the membrane-proximal subdomain, with the formation of multiple hydrogen bonds between the carboxyl and hydroxyl groups of lactate with the side chains of several TlpC residues. ITC experiments on TlpC-LBD variants, demonstrated that only the binding site at the proximal subdomain is involved in lactate recognition by TlpC. These

findings are in stark contrast with recent reports on all other previously characterised dCache chemoreceptors, in which direct ligand recognition has been attributed to the membrane-distal subdomain rather than the proximal subdomain [9,10,17]. Thus, the structural analysis detailed in **Chapter 4** provides the first example of a dCache type chemoreceptor that directly senses its ligand *via* its membrane-proximal subdomain.

Collectively the findings described above showed that chemoreceptors from dCache family can detect their ligands *via* several direct and indirect mechanisms utilising either the membrane-distal, or the membrane-proximal subdomain.

Future work

As described previously, dCache LBDs have a bimodular arrangement with two putative ligand-binding pockets, each of which in principle could directly bind a signal molecule. For example, a 350 $Å^3$ putative ligand-binding pocket is presented at the membrane-proximal subdomain of Tlp3-LBD, which leaves the possibility that additional chemoeffectors could be recognised directly through this subdomain. The sensory specificity of the proximal subdomain in Tlp3-LBD could be explored for example, through screening by molecular docking, identifying novel compounds recognised by this chemoreceptor which may or may not be chemotactically active.

From the branched-chain aliphatic amino acids that interact with Tlp3-LBD, only the chemotactic response triggered by isoleucine has been investigated [11]. Isoleucine is a well-known chemoattractant for *C. jejuni*[11,30]. Therefore, an attractant response could be also expected after the binding of Tlp3-LBD to similar molecules such as leucine, valine and α-amino-N-valeric acid. Analysis of the chemotactic behaviour of wild-type *C. jejuni* and *Δtlp3* mutant strains will be valuable to determine whether the novel Tlp3 ligands identified are chemotactically active and the role of Tlp3 mediating these responses.

As described in **Chapter 3**, *C. jejuni* Tlp1 receptor appears to sense aspartate by an indirect mechanism. However, the PBP that interacts with the dCache LBD of Tlp1 remains to be identified. In *C. jejuni* genome at least 15 different PBPs are annotated, among which, PEB1a, has been identified as an L-aspartate binding protein [31]. PEB1a is the PBP component of an ABC transporter system essential for the intake of aspartate and glutamate by *C. jejuni* [31]. Therefore, it could be hypothesised that PEB1a is involved in aspartate taxis. It will be worth investigating whether Tlp1-LBD interacts with the aspartate binding protein PEB1a or

any of the other 14 *C. jejuni* PBPs to fully elucidate the mechanism by which this amino acid ligand is sensed by Tlp1. Analogously to Tlp1, TlpC-LBD structural analysis also suggests that this dCache domain might recognise ligands through an indirect mechanism i.e., *via* PBP. Therefore, it is also recommended to identify potential PBPs-TlpC receptor interactions to further explore mechanisms for ligand recognition in dCache receptors.

Finally, the work presented in this thesis reduces the knowledge gap regarding dCache sensory domains, leading to the characterisation of new chemoreceptors via mutagenesis, biochemical and structural studies. Furthermore, the derived structural information provides a foundation for future structure-based studies aimed at developing novel therapeutics that target chemoreceptors and the rational redesign of chemoreceptor's specificity which could be used in bioremediation.

References

1 Hendrixson, D. R. & DiRita, V. J. Identification of Campylobacter jejuni genes involved in commensal colonization of the chick gastrointestinal tract. *Mol Microbiol* **52**, 471-484, doi:10.1111/j.1365-2958.2004.03988.x (2004).

2 Terry, K., Williams, S. M., Connolly, L. & Ottemann, K. M. Chemotaxis plays multiple roles during *Helicobacter pylori* animal infection. *Infection and immunity* **73**, 803-811, doi:10.1128/IAI.73.2.803-811.2005 (2005).

3 Dasti, J. I., Tareen, A. M., Lugert, R., Zautner, A. E. & Gross, U. *Campylobacter jejuni*: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. *Int J Med Microbiol* **300**, 205-211, doi:10.1016/j.ijmm.2009.07.002 (2010).

4 Ortega, A., Zhulin, I. B. & Krell, T. Sensory Repertoire of Bacterial Chemoreceptors. *Microbiol Mol Biol Rev* **81**, doi:10.1128/MMBR.00033-17 (2017).

5 Lacal, J., Garcia-Fontana, C., Munoz-Martinez, F., Ramos, J. L. & Krell, T. Sensing of environmental signals: classification of chemoreceptors according to the size of their ligand binding regions. *Environ. Microbiol.* **12**, 2873-2884, doi:10.1111/j.1462-2920.2010.02325.x (2010).

6 Wadhams, G. H. & Armitage, J. P. Making sense of it all: bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.* **5**, 1024-1037, doi:10.1038/nrm1524 (2004).

7 Nishiyama, S. *et al.* Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infection and immunity* **80**, 3170-3178, doi:10.1128/IAI.00039-12 (2012).

8 Nishiyama, S. *et al.* Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci. Rep.* **6**, 20866; 20810.21038/srep20866, doi:10.1038/srep20866 (2016).

9 Glekas, G. D. *et al.* The *Bacillus subtilis* chemoreceptor McpC senses multiple ligands using two discrete mechanisms. *J. Biol. Chem.* **287**, 39412-39418, doi:10.1074/jbc.M112.413518 (2012).

10 Glekas, G. D. *et al.* A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. *J. Biol. Chem.* **285**, 1870-1878, doi:10.1074/jbc.M109.072108 (2010).

11 Rahman, H. *et al.* Characterisation of a multi-ligand binding chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. *PLoS Pathog* **10**, e1003822, doi:10.1371/journal.ppat.1003822 (2014).

12 Chervitz, S. A. & Falke, J. J. Molecular mechanism of transmembrane signaling by the aspartate receptor: a model. *Proc Natl Acad Sci U S A* **93**, 2545-2550 (1996).

13 Yu, D., Ma, X., Tu, Y. & Lai, L. Both piston-like and rotational motions are present in bacterial chemoreceptor signaling. *Sci Rep* **5**, 8640, doi:10.1038/srep08640 (2015).

14 Yeh, J. I. *et al.* High-resolution structures of the ligand binding domain of the wild-type bacterial aspartate receptor. *Journal of molecular biology* **262**, 186-201 (1996).

15 Kim, K. K., Yokota, H. & Kim, S. H. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* **400**, 787-792, doi:10.1038/23512 (1999).

16 Hartley-Tassell, L. E. *et al.* Identification and characterization of the aspartate chemosensory receptor of *Campylobacter jejuni*. *Mol Microbiol* **75**, 710-730, doi:10.1111/j.1365-2958.2009.07010.x (2010).

17 Zhang, Z. & Hendrickson, W. A. Structural characterization of the predominant family of histidine kinase sensor domains. *Journal of molecular biology* **400**, 335-353, doi:10.1016/j.jmb.2010.04.049 (2010).

18 Hazelbauer, G. L. & Adler, J. Role of the galactose binding protein in chemotaxis of Escherichia coli toward galactose. *Nat New Biol* **230**, 101-104 (1971).

19 Adler, J., Hazelbauer, G. L. & Dahl, M. M. Chemotaxis toward sugars in Escherichia coli. *J Bacteriol* **115**, 824-847 (1973).

20 Galloway, D. R. & Furlong, C. E. The role of ribose-binding protein in transport and chemotaxis in Escherichia coli K12. *Arch Biochem Biophys* **184**, 496-504 (1977).

21 Zukin, R. S., Strange, P. G., Heavey, R. & Koshland, D. E. Properties of the galactose binding protein of Salmonella typhimurium and Escherichia coli. *Biochemistry* **16**, 381-386 (1977).

22 Hazelbauer, G. L. Maltose chemoreceptor of Escherichia coli. *J Bacteriol* **122**, 206-214 (1975).

23 Manson, M. D., Blank, V., Brade, G. & Higgins, C. F. Peptide chemotaxis in E. coli involves the Tap signal transducer and the dipeptide permease. *Nature* **321**, 253-256, doi:10.1038/321253a0 (1986).

24 Hegde, M. *et al.* Chemotaxis to the quorum-sensing signal AI-2 requires the Tsr chemoreceptor and the periplasmic LsrB AI-2-binding protein. *J Bacteriol* **193**, 768-773, doi:10.1128/JB.01196-10 (2011).

25 Rico-Jimenez, M. *et al.* Two different mechanisms mediate chemotaxis to inorganic phosphate in Pseudomonas aeruginosa. *Sci Rep* **6**, 28967, doi:10.1038/srep28967 (2016).

26 Machuca, M. A., Liu, Y. C., Beckham, S. A., Gunzburg, M. J. & Roujeinikova, A. The Crystal Structure of the Tandem-PAS Sensing Domain of *Campylobacter jejuni* Chemoreceptor Tlp1 Suggests Indirect Mechanism of Ligand Recognition. *J. Struct. Biol.* **194**, 205-213, doi:10.1016/j.jsb.2016.02.019 (2016).

27 Machuca, M. A. *et al.* Helicobacter pylori chemoreceptor TlpC mediates chemotaxis to lactate. *Sci Rep* **7**, 14089, doi:10.1038/s41598-017-14372-2 (2017).

28 Iwatani, S. *et al.* Identification of the genes that contribute to lactate utilization in *Helicobacter pylori*. *PLoS One* **9**, e103506; 103510.101371/journal.pone.0103506, doi:10.1371/journal.pone.0103506 (2014).

29 Takahashi, T. *et al.* L-lactic acid secreted from gastric mucosal cells enhances growth of *Helicobacter pylori*. *Helicobacter* **12**, 532-540, doi:10.1111/j.1523-5378.2007.00524.x (2007).

30 Hugdahl, M. B., Beery, J. T. & Doyle, M. P. Chemotactic behavior of *Campylobacter jejuni*. *Infection and immunity* **56**, 1560-1566 (1988).

31 Leon-Kempis Mdel, R., Guccione, E., Mulholland, F., Williamson, M. P. & Kelly, D. J. The Campylobacter jejuni PEB1a adhesin is an aspartate/glutamate-binding protein of an ABC transporter essential for microaerobic growth on dicarboxylic amino acids. *Mol Microbiol* **60**, 1262-1275, doi:10.1111/j.1365-2958.2006.05168.x (2006).