Supporting Information

Signal binding at both modules of its dCache domain enables the McpA chemoreceptor of *Bacillus velezensis* to sense different ligands

Authors

Haichao Feng [1], Yu Lv [1], Tino Krell [2], Ruixin Fu [3], Yunpeng Liu [4], Zhihui Xu [1], Wenbin Du [5], Qirong Shen [1], Nan Zhang [1*], Ruifu Zhang [1,4*]

Affiliations

[1] Jiangsu Provincial Key Lab of Solid Organic Waste Utilization, Jiangsu Collaborative Innovation Center of Solid Organic Wastes, Educational Ministry Engineering Center of Resource-Saving Fertilizers, The Key Laboratory of Plant Immunity, Nanjing Agricultural University, Nanjing, Jiangsu, People's Republic of China.

[2] Department of Environmental Protection, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, 18008, Granada, Spain.

[3] School of Biology and Food, Shangqiu Normal University, Shangqiu 476000, People's Republic of China.

[4] Key Laboratory of Microbial Resources Collection and Preservation, Ministry of Agriculture, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, People's Republic of China.

[5] State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, People's Republic of China.

*** Corresponding author**

Ruifu Zhang, College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing, 210095, People's Republic of China. E-mail: rfzhang@njau.edu.cn

Nan Zhang, College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing, 210095, People's Republic of China. E-mail: nanzhang@njau.edu.cn

The authors declare no conflict of interest.

Keywords: signal transduction, chemotaxis, chemoreceptor, dCache sensor domain, ligand recognition

This PDF file includes:

SI Materials and Methods

Figures S1 to S12

Tables S1 to S10

Legend for Dataset S1

SI Materials and Methods

Proteins expression and purification

Protein used for ligand binding assay. The expression strain BL21(DE3) containing the plasmid pET-29a(+)-McpA-LBD(A36-L277) (1) was grown at 37 °C with 220 r.p.m. shaking in 200 mL of LB medium supplemented with 30 mg/L Kan. When the OD_{600} of the culture reached 0.6, the growth temperature was reduced to 15 °C and isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 24 mg/mL. After overnight incubation, cells were harvested by centrifugation at 8,000 g for 10 min, washed twice with PBS (8 g NaCl, 0.2 g KCl, 1.44 g Na_2HPO_4, and 0.24 g KH_2PO_4 in 1 L; pH 7.4), and finally resuspended in 50 mL PBS. Cells were lysed by sonication on ice, and centrifuged at 15,000 g for 30 min at 4 °C. The lysates were passed through 0.22 μm cut-off filters and loaded onto a metal affinity chromatography column (EconoFit Nuvia IMAC Columns, Ni-charged) according to the manufacturer's (Bio-Rad Laboratories, Inc., United Kingdom) instructions. Protein was eluted by applying a 50 to 500 mM imidazole gradient in PBS buffer. Purified proteins were stored in PBS at -80 °C.

The LBDs of the receptor chimeras M1 and M2 were generated as individual purified proteins. The portion of M1-LBD (A36-P278) and M2-LBD (A40-L277) were cloned into a pET28a expression vector to generate an N-terminal 6×His-tagged construct, with a tobacco etch virus (TEV) protease site, under the control of IPTG (1). Proteins were purified as described above.

Protein used for crystallization. The DNA fragment encoding amino acids 38 to 224 of *B. velezensis* SQR9 McpA, corresponding to the LBD, was cloned into the pET28a expression plasmid using restriction enzymes *Nde I* and *BamH I*. The resulting plasmid pET28a-His-TEV-McpA-LBD (Q38-L224), encoding a protein containing an N-terminal 6×His-tag (cleavable by the TEV protease), was transferred into BL21 (DE3). The expression and purification procedures were similar to those described above. Subsequently, the His-tag was removed by incubating the protein with TEV protease at a protease: protein ratio of 1:40 (w/w) overnight at 4 ºC.

LC-MS analysis. Ten μL of the purified proteins at 1mg/mL were used for LC-MS analysis (TOF LC/MS Agilent Technologies 6224, Agilent Technologies, Palo Alto, USA). The mobile phase A was 0.1% formic acid in H_2O, the mobile phase B was 0.1% formic acid in acetonitrile. The molecular weights of the analyzed proteins were precisely detected by LC/MS (TOF LC/MS Agilent Technologies 6224, Agilent Technologies, Palo Alto, USA).

Circular dichroism (CD) analysis

CD spectroscopy was used to compare the secondary structure composition of McpA-LBD, its single-point variants and the LBDs of receptor chimeras. Protein samples were prepared at a concentration of 0.1 mg/mL in PBS (pH 7.4). Far-UV CD spectra were recorded over the wavelength 200-260 nm at 20 °C with the scan rate of 100 nm/min using a JASCO J-1500 spectropolarimeter. The spectra were measured in triplicate, averaged and smoothed using the Savitzky-Golay algorithm (2).

Ligand binding assay

Surface plasmon resonance (SPR). SPR was performed on a SensiQ instrument (The Pioneer platform, ForteBio). McpA-LBD (A36-L277) was immobilized onto the SPR chips (COOH 5, ForteBio), and the binding affinities were determined for 21 McpA ligands that were identified previously (1). Briefly, the chip was washed with 10 mM NaOH and activated with an EDC-NHS solution containing 0.4 M 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC) and 0.1 M N-hydroxysuccinimide (NHS). The protein was diluted in sodium acetate buffer (pH 4.0) at a final concentration of 50 μg/mL, prior to adding 1 M ethanolamine (pH 8.0) to seal the channel. By using the OneStep injection mode, the protein-containing chips were used to monitor the binding of compounds that were prepared at 200 μM and 2 mM in 0.01 M PBS. To measure the binding activity, the differences of ligand binding to protein-containing and protein-free chips were determined. Between each assay, a series of buffer injections was made to enable double-blank subtraction for the sensorgram. The datasets were analyzed using the Pioneer 4 (version 2.1.5) software. The binding signal was quantified as response units (RU).

M1-LBD and M2-LBD were immobilized on a CM7 series S sensor chip, to determine the binding constants of 15 compounds (as shown in **Figure 3**) using a BIAcore T200 instrument (Cytiva). Proteins were captured using amine coupling kit (GE Healthcare), in which the chip was activated with an EDC-NHS solution. Proteins were diluted in 10 mM sodium acetate buffer pH 4.0 at a final concentration of 50 μg/mL, prior to adding 1 M ethanolamine (pH 8.0) to seal the channel. Under these conditions, 28800 and 29000 RU of M1-LBD and M2-LBD were captured on flow cells 2, and 3, respectively. As a negative control, flow cells 1 was a blank control, which underwent the same treatment as the other flow cells 2 and 3, but without the protein immobilization. Compounds were prepared at 200 μM and 2 mM in 1× PBS-P (0.02 M phosphate buffer, 2.7 mM KCl, 137 mM NaCl, 0.05% surfactant P20; pH 7.4). The

differences of ligand binding to protein-containing and protein-free chips were determined. Between each assay, a series of buffer injections were made to enable double-blank subtraction for the sensorgram. The datasets were analyzed using the BIAcore T200 evaluation software set v3.x. The binding signal was quantified as response units (RU).

Biolayer interferometry analysis (BLI). Binding of the 21 ligands to McpA-LBD (A36-L277) was also monitored using the Octet-RED 96 biolayer interferometry device (ForteBio, Inc., Menlo Park, US) at 25 °C. Ligands were dissolved in PBS (pH 7.0) at 1 mM and subsequently diluted to 3.13 µM, 6.25 µM, 12.50 µM 25.00 µM, 50.00 µM, 75.00 µM and 100.00 µM. Super Streptavidin (SSA) biosensors (ForteBio) were used to capture 50 µM biotin-labeled McpA-LBD(A36-L277) onto the surface of the SSA biosensor. Initially, a baseline was recorded using PBS for 60 s, which was followed by protein immobilization on the SSA biosensor tips by agitation for 120 s and subsequent equilibration with PBS for 60 s. Subsequently, the binding of ligands at different concentrations to McpA-LBD(A36-L277) was recorded for 60 s followed by monitoring ligand dissociation using PBS for another 60 s. The BLI data for each binding event were summarized as a "nm shift" (the wavelength/spectral shift in nanometers) and K_D values determined by fitting to a 1:1 binding model. The binding profile of each sample provided as a "nm shift" (the wavelength/spectral shift in nanometers), which represented the association and disassociation of the protein and the ligands analysed.

X-ray crystallography, data collection, and structure determination

Since the full-length McpA-LBD was proteolytically cleaved between amino acids 224 to 225 (***SI Appendix*, Figure S12**), hampering its crystallization (***SI Appendix*, Table S3**), the domain was shortened to amino acids 38 to 224 (McpA-sLBD).

McpA-sLBD in 50 mM Tris-HCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP), 150 mM NaCl, 1 mM phenylmethylsulfonyl fluorid (PMSF), 5 % (v/v) glycerol at a concentration of 10.62 mg/mL was used. The protein was incubated with 5 mM malic acid (pH 7.6) for 2 h on ice, and then centrifuged at 12,000 g and 4 °C for 5 min. The resulting supernatant was used for sitting-drop crystallization trials in which 0.2 µL protein and 0.02 µL seed solution were added to 0.18 µL reservoir solution, and then equilibrated against 15 µL reservoir solution (0.17 M ammonium acetate, 0.085 M sodium acetate pH 4.6, 25.5 % (w/v) PEG 4000, 15 % glycerol). Crystallisation set-ups were kept at 20 °C and crystals appeared after about thirty days. The crystals were directly transferred into the protectant solution using a loop (0.17 M ammonium acetate, 0.085 M sodium acetate pH 4.6, 25.5 % (w/v) PEG 4000, 25 % glycerol) and, following

incubation for 3-5 seconds, incubation were flash-frozen in liquid nitrogen. Crystals were of approximately 600 μm × 90 μm. X-ray diffraction data were collected at the BL17U1 beam lines of the SSRF (Shanghai Synchrotron Radiation Facility, Shanghai, China).

Diffraction data was processed with HKL2000 (3). Proteins were modeled into the electron density using Coot (4), and the model was refined using REFMAC (5). The crystal structure of McpA-sLBD in complex with malic acid was solved by molecular replacement using the structure of PDB entry 6pxy as template and CCP4 (6). Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 7w0w (McpA-sLBD).

In silico analysis

To screen key residues that comprise the ligand binding pocket and govern the ligand specificity of McpA-LBD, a three-dimensional model of the protein (A36-L277) was constructed using Phyre^2 (7) and template PDB ID 3C8C. Briefly, we firstly collected McpA-LBD homologs with known 3D structure using the following steps: (i) dCache domains homologous (amino acids similarity ≥ 25%) to McpA-LBD were retrieved from the NCBI's Protein Data Bank protein database (***SI Appendix*, Table S4**); (ii) the dCache domains that sense the same ligands as McpA, such as the PctA-LBD and Tlp3(CcmL)-LBD in complex with Ile (5T65, and 4XMR, respectively) (8, 9) (10, 11) were selected (iii) structures in which the membrane-proximal module of dCache domains bind ligands, such as TlpC-LBD with lactate (5WBF) (12) and PctA-LBD with acetate (5T65, note that PctA-LBD bound acetate with very low affinity) (9), were selected. We totally retrieved 11 ligand complexed structures of McpA-LBD homologs (PDB codes: 6pxy, 5ltx, 5ltd, 5ltv, 3c8c, 5ave, 6iov, 6iop, 4xmr, 5t65, and 5wbf). Subsequently, the three-dimensional model of McpA-LBD was compared to the ligand co-crystal structures, to identify residues that are within four Å distance of the bound ligand. The identified amino acids are provided in the **Table S4 (*SI Appendix*)**.

Plant material and growth condition

Cucumber seeds of cultivar 'Jinchun 4' were surface-sterilized with 2 % (w/v) NaClO for 15 min, submerged into 70 % (v/v) ethanol for 1 min, washed three times in sterile distilled water, and then placed onto the sterile Petri plates containing sterile filter paper moistened with sterile distilled water. Plates were left in the dark for approximately three days for pregermination, and then placed into sterile tissue culture bottles containing vermiculite (Xianlin Garden Center, Nanjing). Bottles were incubated in a growth chamber at 22 °C for five days with a photoperiod of 16 h light/ 8 h dark. Each germinated seedling was subsequently transplanted into 50 mL

conical flasks containing 40 mL sterile 1/4 sucrose-free Murashige Skoog (MS) medium (13). The medium was replaced every second day during the growth to minimize potential contaminations. Subsequently, the seedling containing flasks were cultured in a shaker at 50 r.p.m. and exposed to cool white fluorescent light with a photoperiod of 16 h light/8 h dark at 25 ± 5°C. The cucumber plants with fully developed cotyledons were used for further studies. Prior to the conduct of the root colonization assays, the sterility of the MS medium was verified by plating 100-µL aliquots onto LLB agar plates.

Root colonization assay

Suspensions of C-*mcpA* and mutant strains were prepared in LLB medium (containing 5 mg/L Cm if necessary), and grown to the mid-log phase (OD_{600} 1.0) at 30 °C, collected, and then resuspended in 1/4 MS medium. These cells were used to inoculate at an OD_{600} of 0.01 the sterile conical flasks containing 50 mL MS medium and the growing cucumber plants. After three days, the cucumber roots were harvested and briefly washed three times with sterile distilled water. Root samples were weighed and ground with a mortar in 10 mL PBS. Serial dilutions were plated onto LLB agar medium containing 5 mg/L cm. After an incubation at 30 °C for 12 h the colonies were counted and normalized with the root fresh weight. Six replicates for each treatment were analysed.

Conservation analysis

To further uncover the potential conservation of amino acid residues in dCache domains involved in sensing of different categories of ligands (amino acids, organic acids and sugars/sugar alcohols), the sequences of chemoreceptors that were reported to respond to ligands of these three ligand categories were retrieved from NCBI database. We initially compiled the dCache chemoreceptors that bind ligands of these three categories based on the following criteria: (i) A simple conserved motif has been recently identified in the dCache (9, 14), which provides a common molecular mechanism for amino acid sensing. The conserved sequence motif for amino acid recognition was Tyr121, Arg126, Trp128, Tyr129, Tyr144, and Asp173 (according to PctA numbering); (ii) The TlpA chemoreceptor of *H. pylori* SS1 was reported to bind fumarate involving residues Asp165, Met183, Tyr228, Tyr252 and Asp254 (15), while TlpC in *H. pylori* 26695 bound lactate through interactions with primarily Asn213, Ile218, Tyr249 and Tyr285 (12). Moreover, residues Thr199, Gln200, His273 and Glu274 were associated in the McpA chemoreceptor of *B. subtilis* OI1085 with sensing acidic pH (16). The dCache McpH of *P. putida* KT2440 could bind uric acid, but the amino acid residues involved

are unknown (17). However, it should be noted that the majority of chemoreceptors that respond to organic acids mainly employ a LBD of sCache (10, 18, 19), 4HB (20–23), or HBM (24–26) families; (iii) As far as we know, Tlp11 of *C. jejuni* is the only dCache chemoreceptor that can directly bind galactose (27). The retrieved sequences were aligned using MAFFT (v7.475) with default parameter values (28), and the aligned sequences were used for further analyses. Conservation, consensus analysis and visualization of alignment sequences were performed using the CLC Sequence Viewer (v8.0).

Identification of homologs and phylogenetic analyses

The sequences of McpA-LBD were used for BLAST searches in the NCBI database (16 April 2022). Only the hits with coverage higher than 98% and sequence identity superior to 30% were retained for further analyses, resulting in a data set of 4669 sequences that belong to 46 genera and 218 species (**Dataset S1**). Multiple sequence alignments using full-length protein sequences were made using MAFFT (v7.475) with default parameters (28). A phylogenetic tree was constructed using FastTree v2.1.10 software with the LG substitution model (29). The resulting tree was subsequently visualized and edited with iTOL v6.5.2 (https://itol.embl.de/).

SI References

1. H. Feng, *et al.*, Identification of chemotaxis compounds in root exudates and their sensing chemoreceptors in plant growth-promoting rhizobacteria *Bacillus amyloliquefaciens* SQR9. *Mol. Plant-Microbe Interact.* **31**, 995–1005 (2018).
2. N. J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2006).
3. Wladek, *et al.*, HKL-3000: the integration of data reduction and structure solution from diffraction images to an initial model in minutes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 859–866 (2006).
4. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
5. G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255 (1997).
6. S. M. Bailey, The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763 (1994).
7. L. A. Kelley, M. J. E. Sternberg, Protein structure prediction on the web: a case study

using the phyre server. *Nat. Protoc.* **4**, 363–373 (2009).

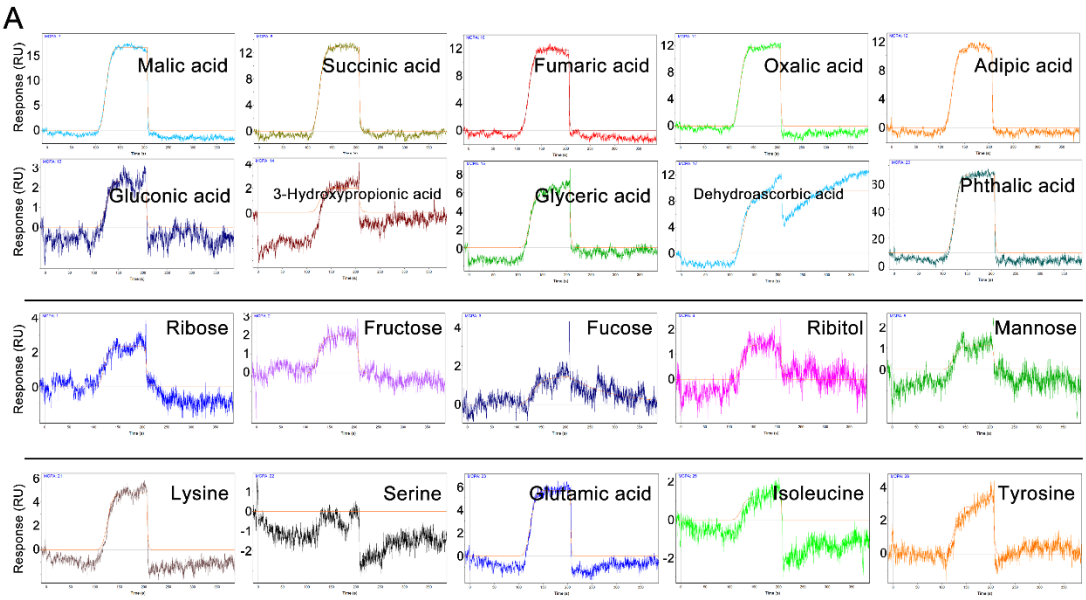
8. M. Rico-Jiménez, *et al.*, Paralogous chemoreceptors mediate chemotaxis towards protein amino acids and the non-protein amino acid gamma-aminobutyrate (GABA). *Mol. Microbiol.* **88**, 1230–1243 (2013).
9. J. A. Gavira, *et al.*, How bacterial chemoreceptors evolve novel ligand specificities. *MBio* **11**, e03066-19 (2020).
10. H. Rahman, *et al.*, Characterisation of a multi-ligand binding chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. *PLoS Pathog* **10**, e1003822 (2014).
11. Y. C. Liu, M. A. Machuca, S. A. Beckham, M. J. Gunzburg, A. Roujeinikova, Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 2127–2136 (2015).
12. M. A. Machuca, *et al.*, *Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate. *Sci Rep* **7**, 14089 (2017).
13. T. Murashige, F. Skoog, A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497 (2006).
14. V. M. Gumerov, *et al.*, Amino acid sensor conserved from bacteria to humans. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2110415119 (2022).
15. K. S. Johnson, *et al.*, The dCache chemoreceptor TlpA of *Helicobacter pylori* binds multiple attractant and antagonistic ligands via distinct sites. *MBio* **12**, e01819-21 (2021).
16. P. Tohidifar, M. J. Plutz, G. W. Ordal, C. V. Rao, The mechanism of bidirectional pH taxis in *Bacillus subtilis*. *J. Bacteriol.* **202**, e00491-19 (2020).
17. M. Fernández, B. Morel, A. Corral-Lugo, T. Krell, Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol. Microbiol.* **99**, 34–42 (2016).
18. V. Garcia, *et al.*, Identification of a chemoreceptor for C2 and C3 carboxylic acids. *Appl Env. Microbiol* **81**, 5449–5457 (2015).
19. K. Karl Compton, S. B. Hildreth, R. F. Helm, B. E. Scharf, *Sinorhizobium meliloti* chemoreceptor McpV senses short-chain carboxylates via direct binding. *J. Bacteriol.* **200**, 1–16 (2018).
20. B. Ni, Z. Huang, Z. Fan, C. Y. Jiang, S. J. Liu, *Comamonas testosteroni* uses a chemoreceptor for tricarboxylic acid cycle intermediates to trigger chemotactic responses towards aromatic compounds. *Mol Microbiol* **90**, 813–823 (2013).
21. Z. Huang, *et al.*, Direct sensing and signal transduction during bacterial chemotaxis toward

aromatic compounds in *Comamonas testosteroni*. *Mol. Microbiol.* **101**, 224–237 (2016).

22. R. A. Luu, *et al.*, Integration of chemotaxis, transport and catabolism in *Pseudomonas putida* and identification of the aromatic acid chemoreceptor PcaY. *Mol. Microbiol.* **96**, 134–147 (2015).
23. T. Iwama, *et al.*, Differential recognition of citrate and a metal-citrate complex by the bacterial chemoreceptor Tcp. *J. Biol. Chem.* **281**, 17727–17735 (2006).
24. E. Pineda-Molina, *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 (2012).
25. J. Lacal, *et al.*, Identification of a chemoreceptor for tricarboxylic acid cycle intermediates. *J. Biol. Chem.* **285**, 23126–23136 (2010).
26. D. Martin-Mora, *et al.*, McpQ is a specific citrate chemoreceptor that responds preferentially to citrate/metal ion complexes. *Env. Microbiol* **18**, 3284–3295 (2016).
27. C. J. Day, *et al.*, A direct-sensing galactose chemoreceptor recently evolved in invasive strains of *Campylobacter jejuni*. *Nat. Commun.* **7**, 13206 (2016).
28. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
29. S. Q. Le, O. Gascuel, An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008).

Figures S1 to S12

A



B

Name	K_a (M^{-1} s^{-1})	K_d (s^{-1})
Malic acid	367.56	0.75
Succinic acid	41.53	0.60
Fumaric acid	92.61	0.82
Oxalic acid	179.18	0.73
Adipic acid	51.01	0.81
Gluconic acid	24.20	0.59
3-Hydroxypropionic acid	6.25	0.36
Glyceric acid	13.03	0.45
Dehydroascorbic acid	44.03	0.00
Phthalic acid	34.99	0.61

Name	K_a (M^{-1} s^{-1})	K_d (s^{-1})
Ribose	173.20	0.26
Fructose	13.60	0.27
Fucose	25.05	0.01
Ribitol	26.44	0.21
Mannose	22.54	0.29

Name	K_a (M^{-1} s^{-1})	K_d (s^{-1})
Lysine	10.62	0.58
Serine	12.44	1.29
Glutamic acid	37.58	0.87
Isoleucine	82.86	1.83
Tyrosine	246.60	0.47

Figure S1. Ligand binding at McpA-LBD by surface plasmon resonance (SPR). (A) Binding response of each ligand to McpA-LBD (A36-L277). The interaction was quantified as response units (RU). One RU equals 1×10^{-6} of refractive index unit, indicating 1 pg of bound molecules per square millimeter of the protein containing sensor chip (B) Kinetic binding parameters as calculated by Pioneer 4 (v.2.1.5).

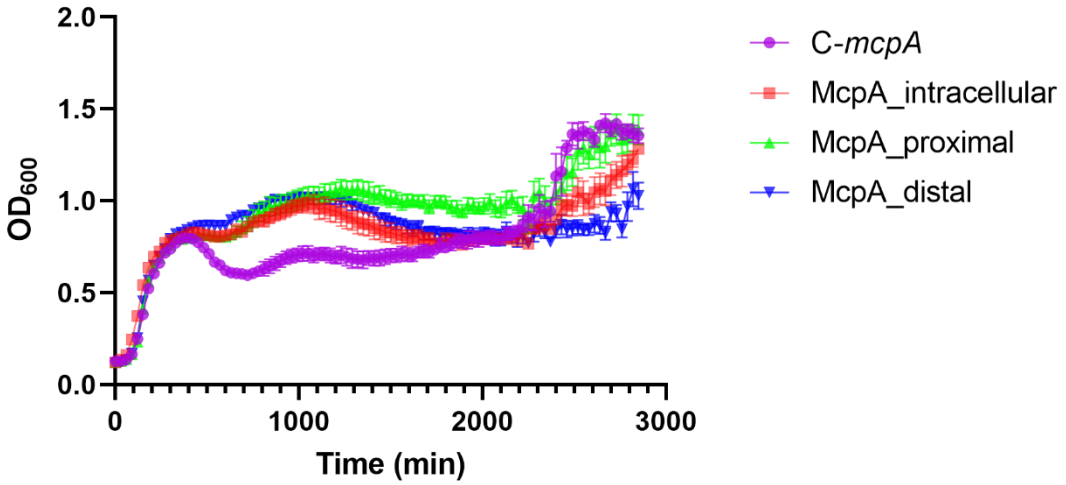


Figure S2. Growth curves of the complementary strain C-*mcpA*, and the three strains containing McpA chimera. Overnight cultures of strains in LLB were adjusted to an OD_{600} =1.0, and the resulting suspension was used to inoculate fresh LLB broth at ratio of 1 % (v/v).

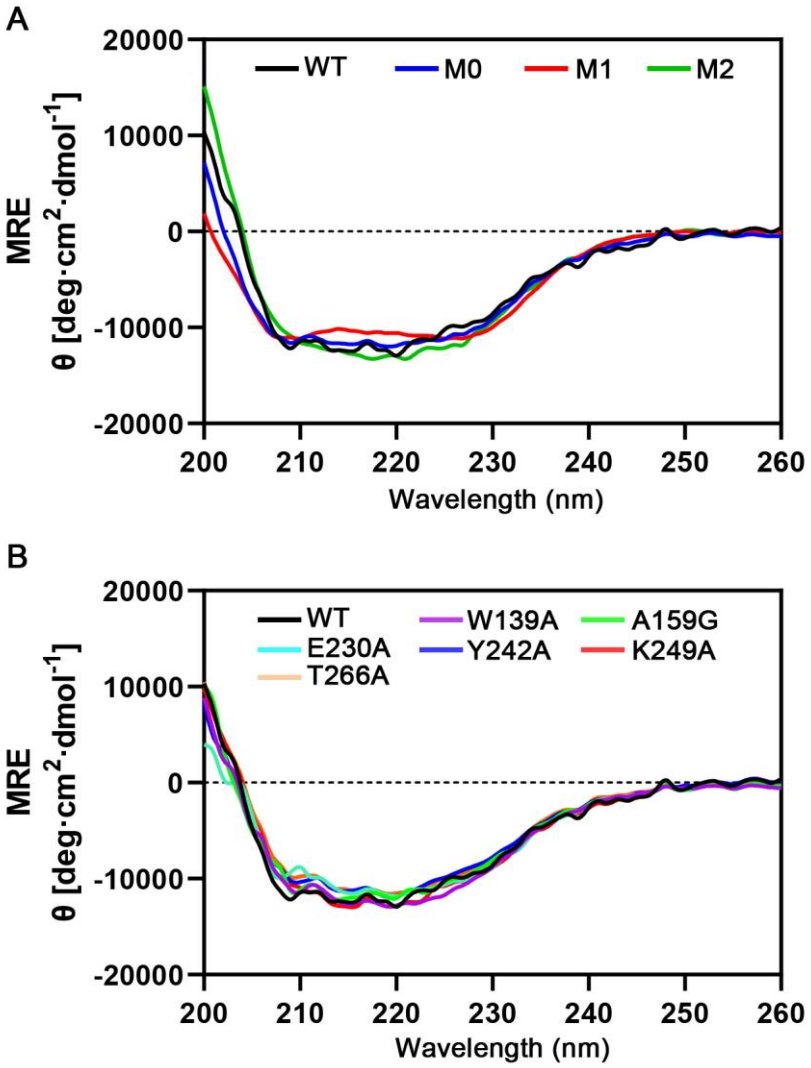
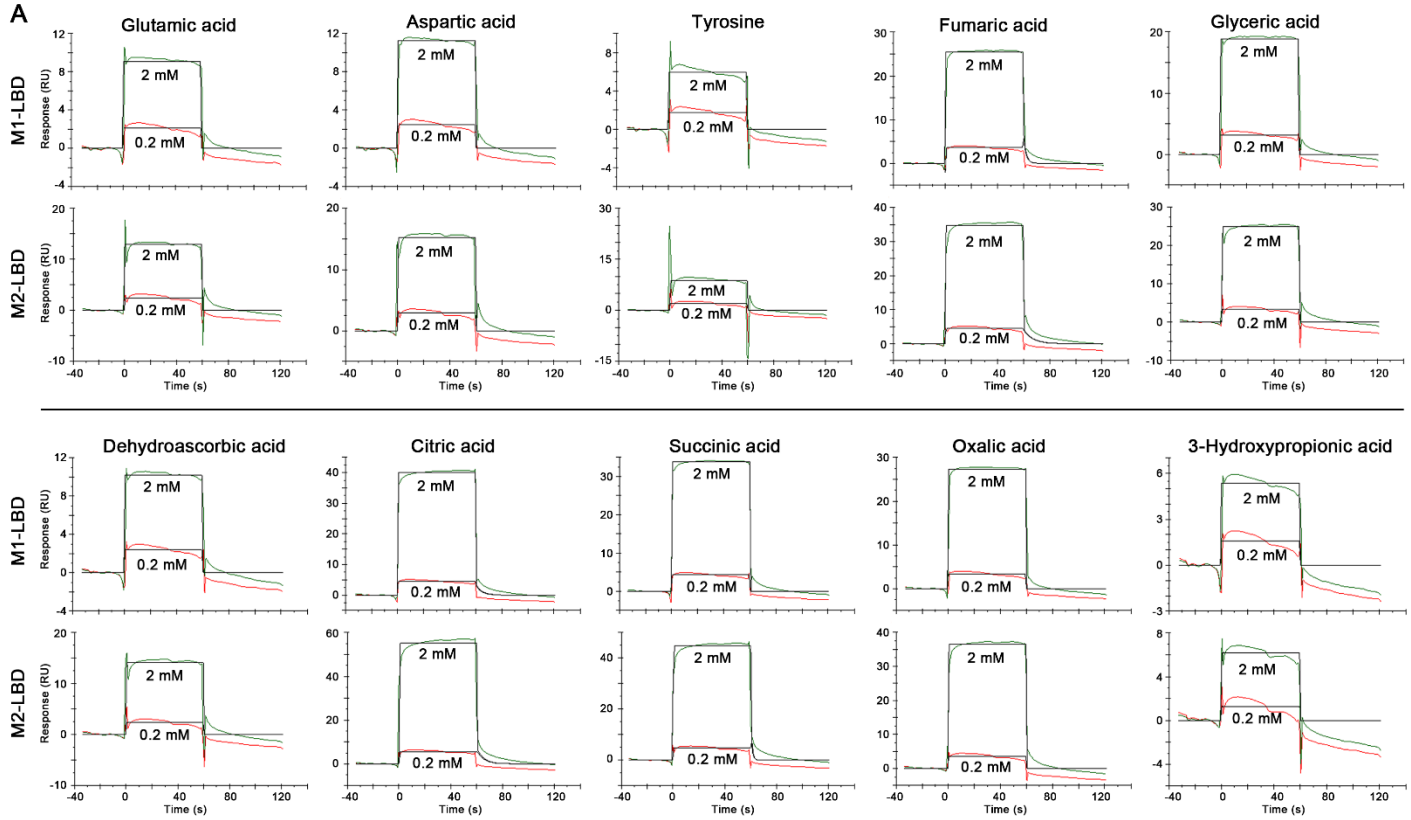


Figure S3. CD spectra of native McpA-LBD (WT), three receptor chimeras (A; M0, M1, and M2) and the six variants (B; W139A, A159G, E230A, Y242A, K249A, and T266A).

A



B

Name	M1-LBD		M2-LBD	
	K_a (M^{-1} s^{-1})	K_d (s^{-1})	K_a (M^{-1} s^{-1})	K_d (s^{-1})
Glutamic acid	1.13×10^{5}	0.004	9.06×10^{4}	0.004
Aspartic acid	1.08×10^{5}	0.004	8.39×10^{4}	0.005
Tyrosine	1.00×10^{5}	0.002	1.41×10^{5}	0.001
Fumaric acid	2.38×10^{7}	0.523	2.07×10^{4}	0.142
Glyceric acid	2.75×10^{4}	0.003	6.21×10^{4}	0.004
Dehydroascorbic acid	1.17×10^{5}	0.005	8.82×10^{4}	0.005
Citric acid	7.77×10^{6}	0.217	1.70×10^{7}	0.187
Succinic acid	7.42×10^{4}	0.025	1.05×10^{8}	0.955
Oxalic acid	9.96×10^{4}	0.003	8.94×10^{4}	0.004
3-Hydroxypropionic acid	9.73×10^{4}	0.002	1.12×10^{5}	0.001

Figure S4. Interaction between the ligand binding domains of receptor chimeras and 15 ligands. (A) Shown are SPR data for the interaction of M1-LBD (A36-P278) and M2-LBD (A40-L277) with members of the three ligand categories. The interaction was quantified as response units (RU). (B) Kinetic binding parameters as calculated by BIAcore T200 evaluation software set v3.x.

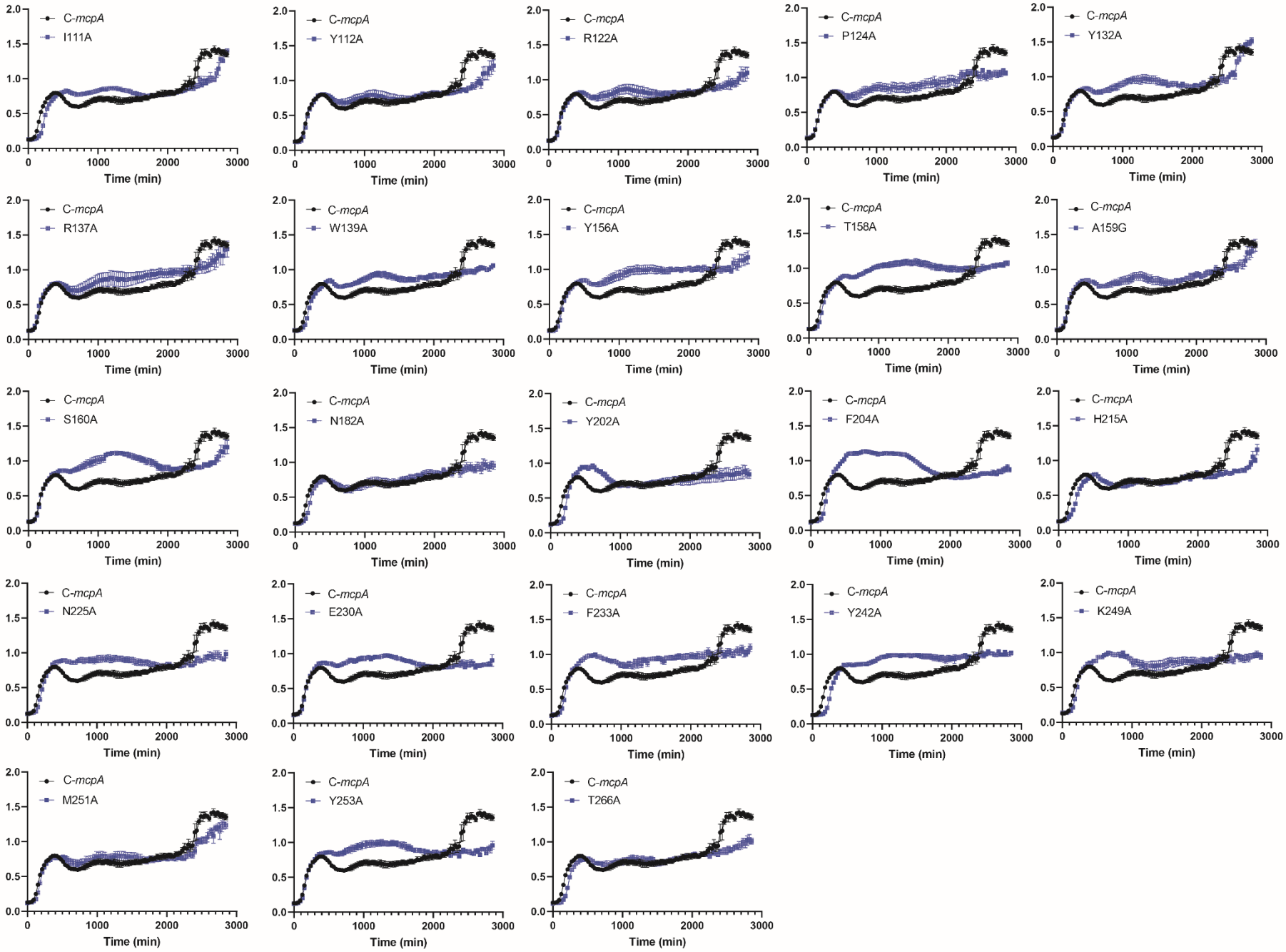

C-mcpA
I111A
Y112A
R122A
P124A
Y132A
R137A
W139A
Y156A
T158A
A159G
S160A
N182A
Y202A
F204A
H215A
N225A
E230A
F233A
Y242A
K249A
M251A
Y253A
T266A
Time (min)

Figure S5. Growth curves of the complementary strain C-*mcpA* and the twenty-three strains harboring McpA point-mutations. Overnight cultures of strains in LLB were adjusted to an OD_{600} =1.0, and the resulting suspension was used to inoculate fresh LLB broth at a ratio of 1 %.

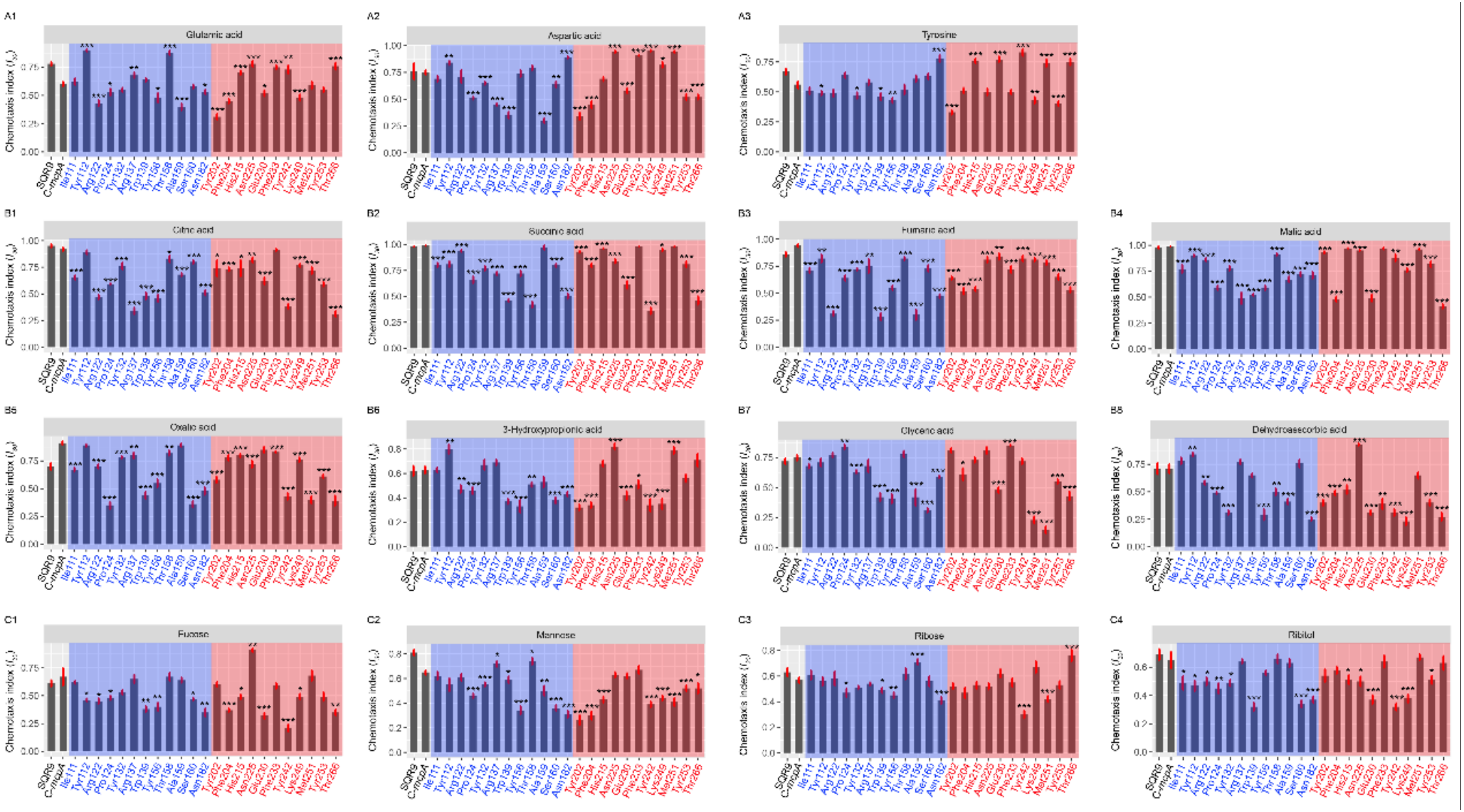


Figure S6. Results of chemotaxis assays of 23 strains harboring point mutations in the McpA-LBD to 15 ligands, including amino acids (A1-A3), organic acids (B1-B8), and sugars/sugar alcohols (C1-C4). The chemotaxis reference data of SQR9 and its complementary strain C-*mcpA* have been published previously (Feng *et al*., 2018). Asterisks indicate statistically significant differences compared with the complementary strain C-*mcpA* (according to the Duncan's multiple rang tests). I_{30} is the chemotaxis index, $I_{30} >$ 0.6 or < 0.4 indicates attractant or repellent responses, respectively; while $0.4 \leq I_{30} \leq 0.6$ corresponds to an absence of taxis. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Data are

means and standard deviations from 13 independent measurements.

Reference

H. Feng, et al., Identification of chemotaxis compounds in root exudates and their sensing chemoreceptors in plant growth-promoting rhizobacteria *Bacillus amyloliquefaciens* SQR9. Mol. Plant-Microbe Interact. 31, 995–1005 (2018).

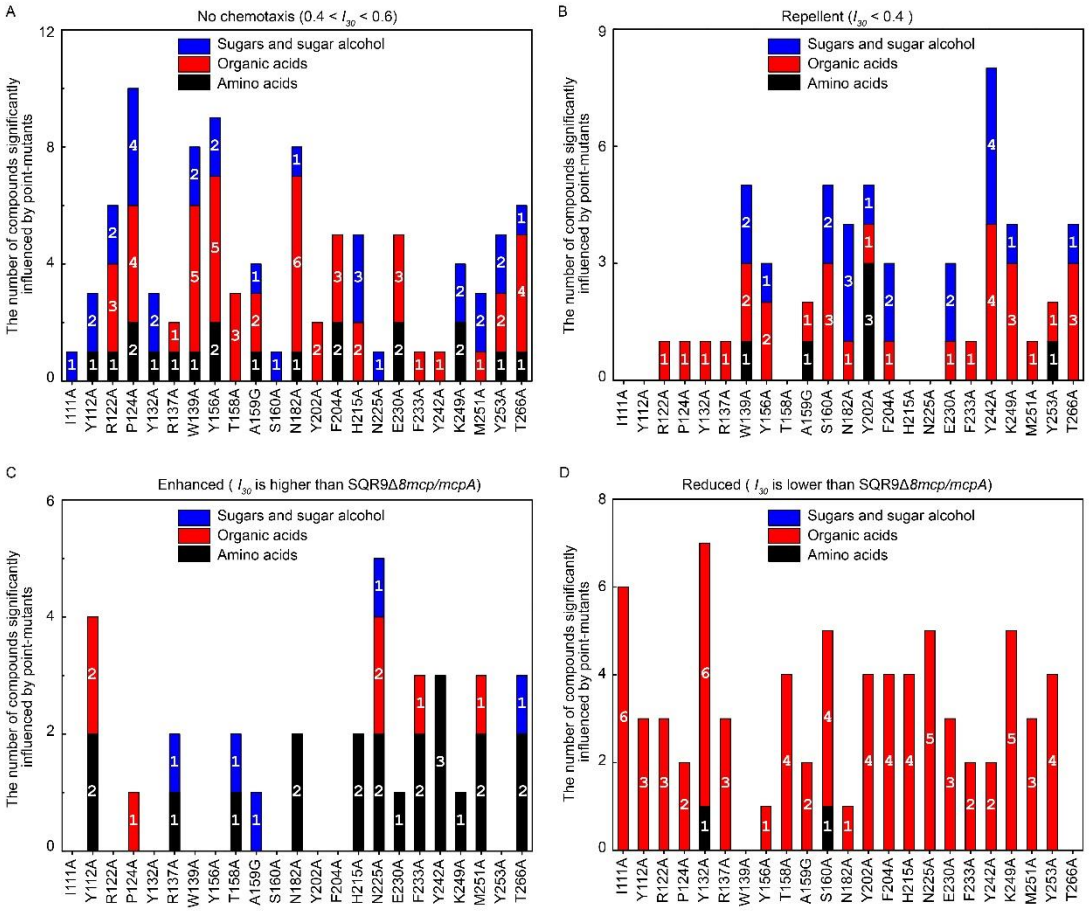


Figure S7. Summary of the impact of the individual point mutations in McpA-LBD on the chemotactic behavior of *B. velezensis* SQR9, including no chemotaxis (A), repellent- (B), enhanced- (C), and reduced- responses (D).

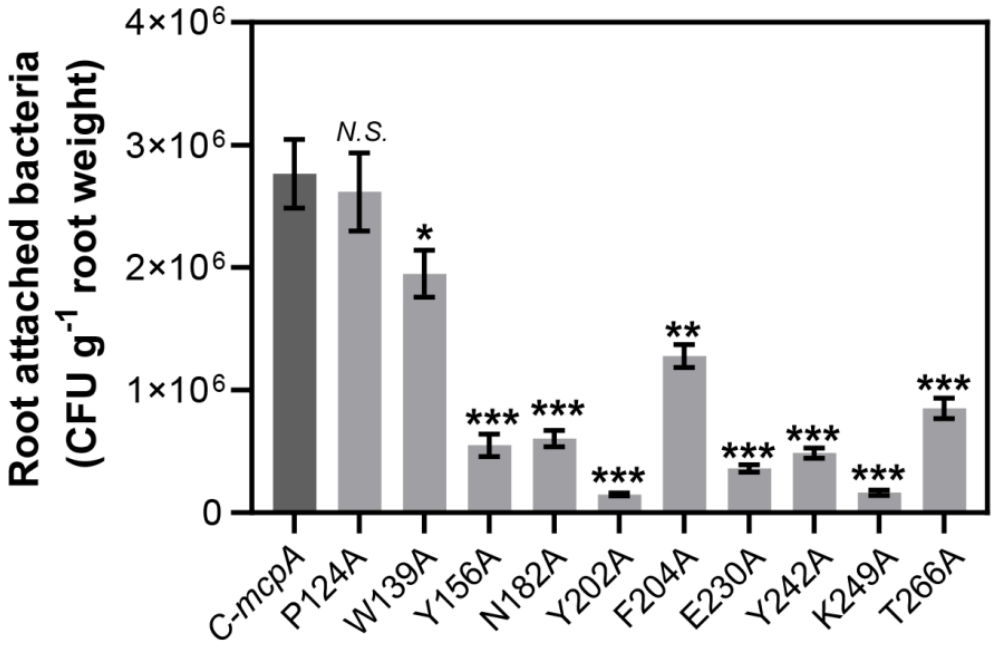


Figure S8. Root colonization capacity of reference strain C-*mcpA* and ten mutant strains harboring mutations in McpA-LBD. * $P < 0.05$, ** $P < 0.01$, indicates $P < 0.001$, *N.S.* indicates non-significance. Data are means and standard deviations from six independent experiments.

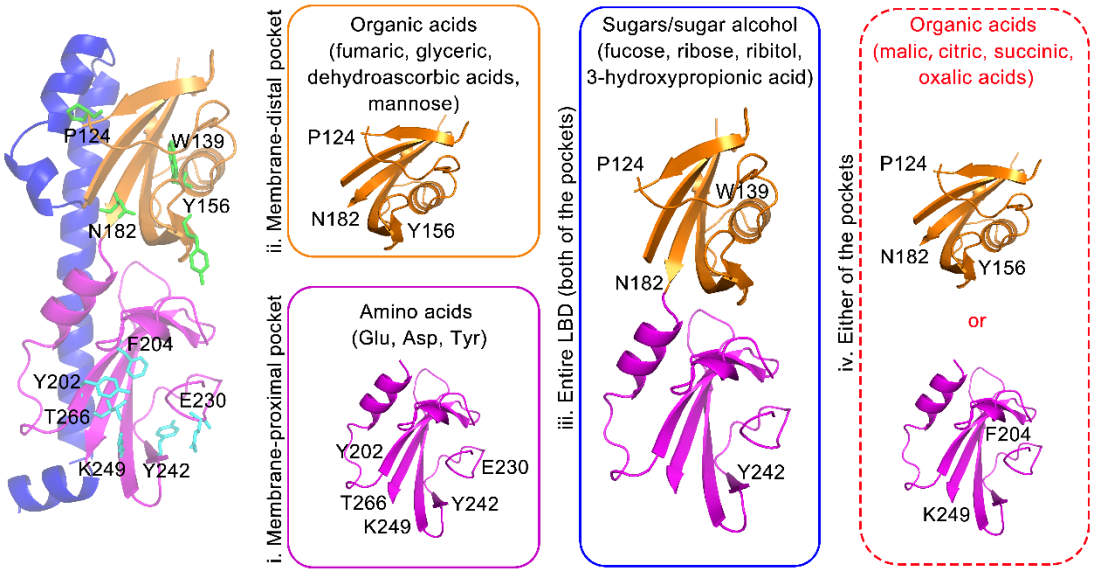
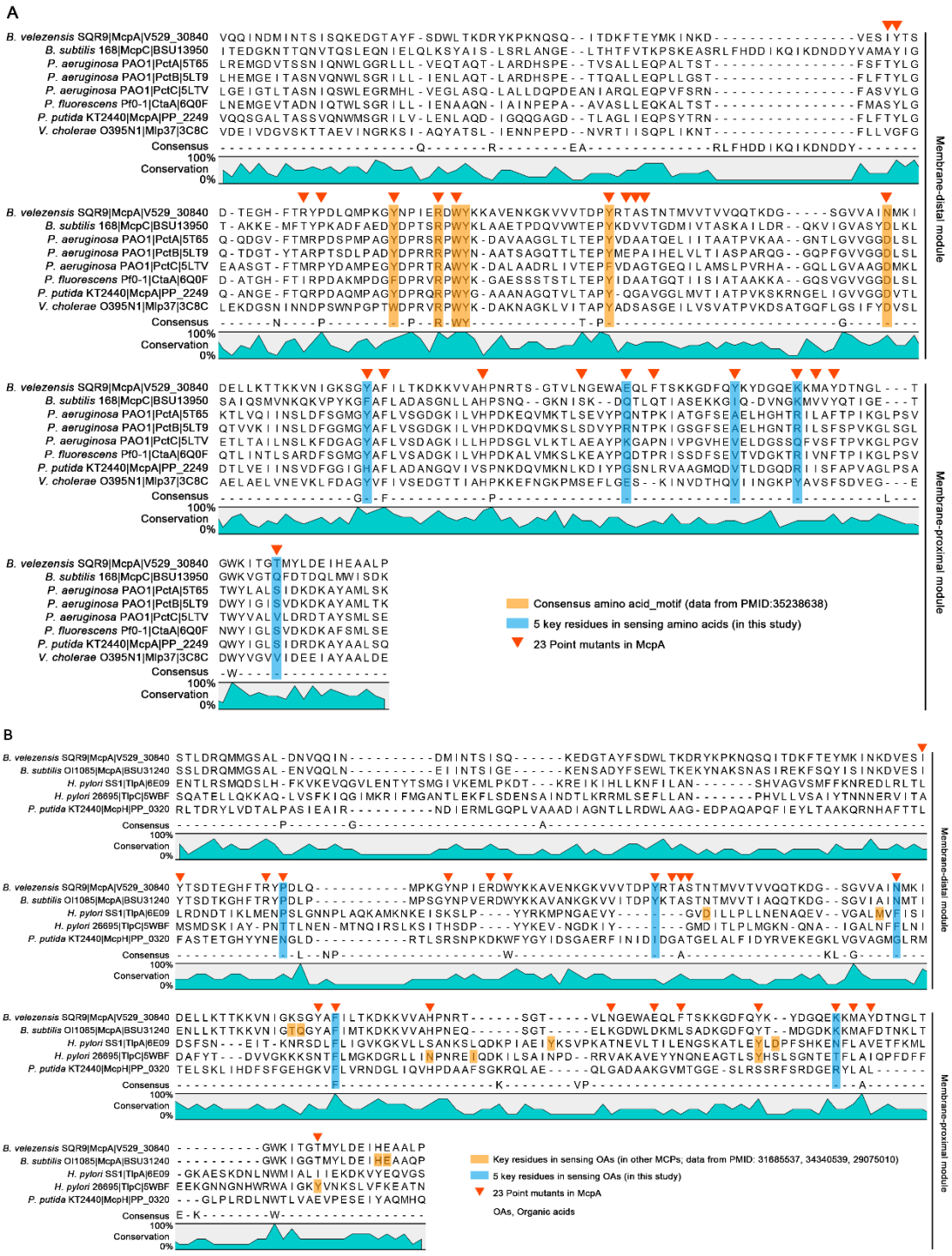


Figure S9. Model for the multi-ligand sensing capacity of McpA. The membrane-distal and membrane-proximal modules of dCache domains are shown in orange and magenta, respectively. The 10 key amino acids are shown in stick mode with carbons colored green (4 in membrane-distal) and cyan (6 in membrane-proximal), respectively. Sensing is hypothesized to occur in 4 different modes. **Mode i** (magenta box): the membrane-proximal module of McpA-LBD specifically senses AAs, namely glutamic acid, aspartic acid, and tyrosine; **Mode ii** (orange box): the membrane-distal domain of McpA-LBD recognize the 3 OAs (fumaric, glyceric, and dehydroascorbic acids) and mannose; **Mode iii** (blue box): the entire LBD domain (both distal and proximal region; P124, W139, N182, and Y242) is required for sensing 3 SAs (fucose, ribose, and ribitol) and 3-hydroxypropionic acid; **Strategy iv** (red dotted boxes): either the membrane-distal or -proximal domain (P124, Y156, and N182; F204 and K249) mediate the response to the four OAs (malic, citric, succinic and oxalic acids).


A
B. velezensis SQR9|McpA|V529_30840
B. subtilis 168|McpC|BSU13950
P. aeruginosa PAO1|PctA|5T65
P. aeruginosa PAO1|PctB|5LT9
P. aeruginosa PAO1|PctC|5LTV
P. fluorescens Pf0-1|CtaA|6Q0F
P. putida KT2440|McpA|PP_2249
V. cholerae O395N1|Mlp37|3C8C
Consensus
100%
Conservation
0%
Membrane-distal module
Membrane-proximal module
Consensus amino acid_motif (data from PMID:35238638)
5 key residues in sensing amino acids (in this study)
23 Point mutants in McpA
B
B. velezensis SQR9|McpA|V529_30840
B. subtilis OI1085|McpA|BSU31240
H. pylori SS1|TlpA|6E09
H. pylori 26695|TlpC|5WBF
P. putida KT2440|McpH|PP_0320
Key residues in sensing OAs (in other MCPs; data from PMID: 31685537, 34340539, 29075010)
5 key residues in sensing OAs (in this study)
23 Point mutants in McpA
OAs, Organic acids

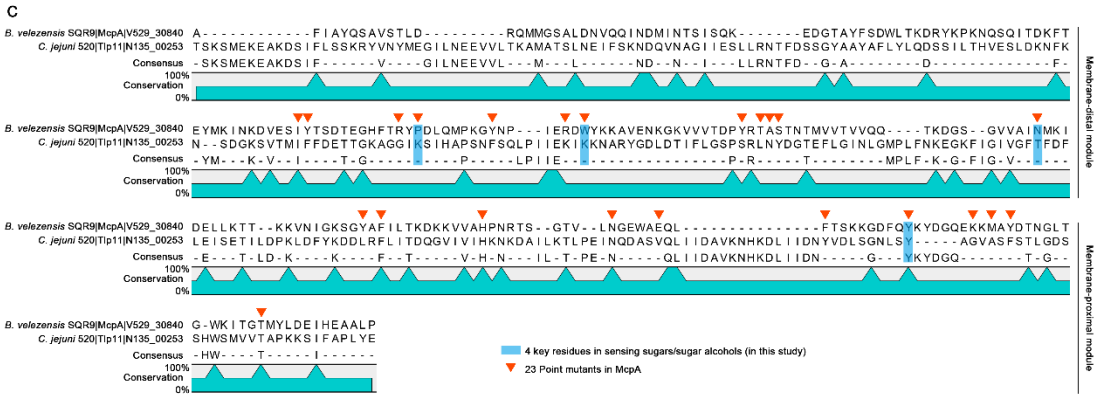


Figure S10. Conservation of the three ligand categories-sensing amino acid residues in dCache domains (A, amino acids; B, organic acids; C, sugars/sugar alcohols). Amino acids sequence alignment of the LBDs of homologous chemoreceptors. The amino acid binding motif (Gumerov *et al*., 2022) is shaded in orange, the amino acid that were required for amino acid sensing are shaded in blue, the asterisks indicate the positions of the 23 single amino acid replacements in McpA. NCBI accession numbers or the PDB IDs are indicated for each protein.

Reference

V. M. Gumerov, et al., Amino acid sensor conserved from bacteria to humans. Proc. Natl. Acad. Sci. U. S. A. 119, e2110415119 (2022).

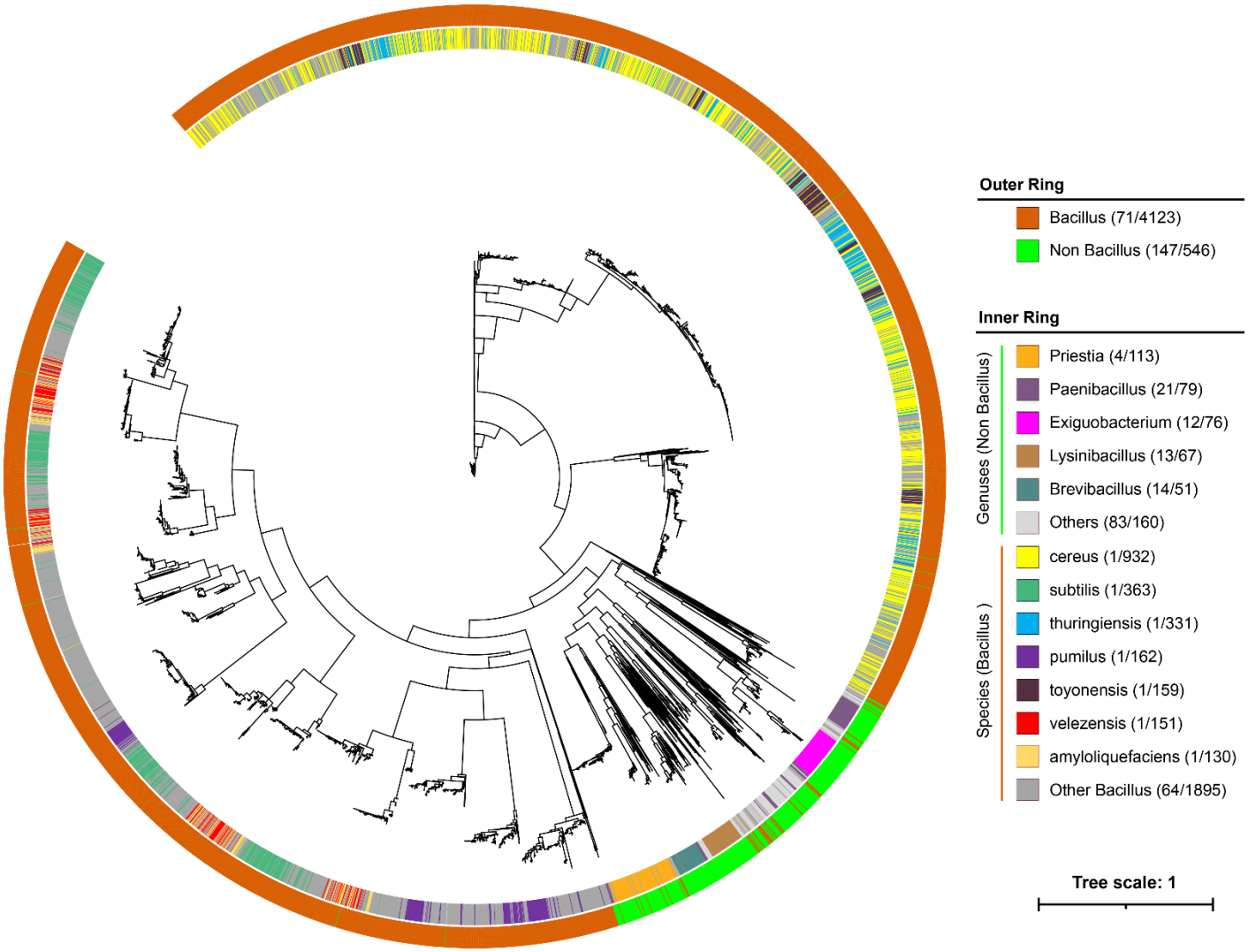


Figure S11. Phylogenetic tree of McpA homologs. These 4669 sequences belong to 46 genera and 218 species, of which only the top six genera and seven species with the highest abundance are shown in the phylogenetic tree. Numbers in parentheses indicate the number of species and sequences contained in the genus, respectively. For detailed information of the species, please refer to the **Dataset S1**.

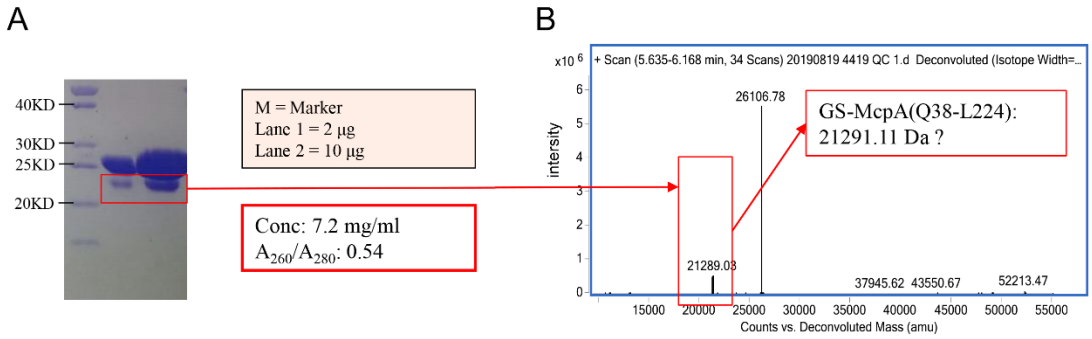


Figure S12. The degradation of McpA-LBD. (A) Analysis of freshly purified protein using the SDS-PAGE. (B) Identification of the cleavage site at McpA-LBD using LC-MS. Shown is the singly charged species in positive ion mode of a sample of the degraded protein that was excised from the SDS-PAGE gel.

Tables S1 to S10

Table S1. Summary of experimental binding studies using SPR, BLI and ITC.

Category	Substances	Chemotaxis assay	Chemotaxis Index (I_{30})		Instruments	K_D (μM)	References (binding parameter)
			SQR9	SQR9Δ*8mcp/mcpA*			
Organic acids	Succinic acid [a]	microfluidic SlipChip device; Chemoattractants	0.98	0.99	SPR and BLI	473±261, BLI	In this study
	Malic acid [a]	microfluidic SlipChip device; Chemoattractants	0.98	0.99	SPR and BLI	4.02±0.24, BLI	In this study
	Gluconic acid [b]	microfluidic SlipChip device; Chemoattractants	0.97	0.66	SPR and BLI	**None, BLI**	In this study
	Citric acid [a]	microfluidic SlipChip device; Chemoattractants	0.95	0.92	ITC	0.39±0.15, ITC	*Feng et al*., 2018
	Fumaric acid [a]	microfluidic SlipChip device; Chemoattractants	0.86	0.94	SPR and BLI	22.2±1.32, BLI	In this study
	Adipic acid [a]	microfluidic SlipChip device; Chemoattractants	0.79	0.65	SPR and BLI	**None, BLI**	In this study
	Phthalic acid [a]	microfluidic SlipChip device; Chemoattractants	0.77	0.58	SPR and BLI	**None, BLI**	In this study
	Glyceric acid [a]	microfluidic SlipChip device; Chemoattractants	0.72	0.75	SPR and BLI	11.5±1.20, BLI	In this study
	Dehydroascorbic	microfluidic SlipChip	0.71	0.71	SPR and BLI	32.9±3.16, BLI	In this study

	acid [a]	device; Chemoattractants					
	Oxalic acid [a]	microfluidic SlipChip device; Chemoattractants	0.70	0.91	SPR and BLI	2.89±0.21, BLI	In this study
	3-Hydroxypropionic acid [a]	microfluidic SlipChip device; Chemoattractants	0.62	0.63	SPR and BLI	58.5±6.41, BLI	In this study
Amino acids	Serine [b]	microfluidic SlipChip device; Chemoattractants	0.99	0.68	SPR and BLI	**None, BLI**	In this study
	Isoleucine [a]	microfluidic SlipChip device; Chemoattractants	0.90	0.58	SPR and BLI	**None, BLI**	In this study
	Lysine [a]	microfluidic SlipChip device; Chemoattractants	0.81	0.57	SPR and BLI	**None, BLI**	In this study
	Glutamic acid [a]	microfluidic SlipChip device; Chemoattractants	0.78	0.60	SPR and BLI	24.9±2.14, BLI	In this study
	Aspartic acid [a]	microfluidic SlipChip device; Chemoattractants	0.76	0.75	ITC	0.24±0.16, ITC	*Feng et al*., 2018
	Tyrosine [a]	microfluidic SlipChip device; Chemoattractants	0.67	0.56	SPR and BLI	133±32.8, BLI	In this study
Sugars	Galactose [a]	microfluidic SlipChip device; Chemoattractants	0.99	0.98	BLI	**None, BLI**	*Liu et al*., 2020
	Mannose [a]	microfluidic SlipChip device; Chemoattractants	0.81	0.65	SPR and BLI	28.6±2.22, BLI	In this study
	Fructose [b]	microfluidic SlipChip device; Chemoattractants	0.74	0.62	SPR and BLI	**None, BLI**	In this study
	Ribitol [a]	microfluidic SlipChip device; Chemoattractants	0.69	0.65	SPR and BLI	4.36±0.60, BLI	In this study

	Ribose [a]	microfluidic SlipChip device; Chemoattractants	0.63	0.57	SPR and BLI	1.94±0.21, BLI	In this study
	Fucose [a]	microfluidic SlipChip device; Chemoattractants	0.61	0.67	SPR and BLI	63.4±6.42, BLI	In this study

Note: [a] Compounds were dominantly sensed by McpA, which indicating that this chemoreceptor predominantly mediates the chemotaxis of SQR9 to these compounds; [b] McpA was partially involved in chemotaxis to these chemicals. Both serine and gluconic acid were dominantly sensed by McpC, the fructose was dominantly sensed by TlpB.

Table S2. Chemotaxis indices of *B. velezensis* SQR9, reference strain SQR9Δ*8mcp/mcpA* and that of 23 McpA-LBD mutants containing single amino acid substitutions for 15 different chemoattractants.

Strains		Chemotaxis Index (I_{30}) [abc]														
		Glutamic acid	**Aspartic acid**	**Tyrosine**	**Malic acid**	**Citric acid**	**Succinic acid**	**Fumaric acid**	**Oxalic acid**	**3-Hydroxypropionic acid**	**Glyceric acid**	**Dehydroascorbic acid**	**Ribose**	**Fucose**	**Mannose**	**Ribitol**
SQR9		0.78	0.76	0.67	0.98	0.95	0.98	0.86	0.70	0.62	0.72	0.71	0.63	0.61	0.81	0.69
SQR9Δ*8mcp/mcpA*, C-*mcpA*		0.60	0.75	0.56	0.99	0.92	0.99	0.94	0.91	0.63	0.75	0.71	0.57	0.67	0.65	0.65
Chimera mutant strains (3)	***mcpA*_intracellular, M0**	0.54	0.37	0.47	0.58	0.49	0.45	0.51	0.41	0.46	0.50	0.54	0.38	0.52	0.51	0.43
	***mcpA*_distal, M1**	0.56	0.50	0.56	0.95	0.75	0.95	0.80	0.66	0.44	0.79	0.64	0.51	0.44	0.72	0.49
	***mcpA*_proximal, M2**	0.71	0.73	0.71	0.78	0.85	0.75	0.51	0.78	0.58	0.58	0.45	0.40	0.48	0.55	0.58
membrane-distal mutants (12)	**I111A**	0.62	0.69	0.51	0.77	0.65	0.80	0.71	0.67	0.63	0.68	0.78	0.61	0.62	0.62	0.49
	Y112A	0.90	0.84	0.49	0.90	0.89	0.81	0.82	0.89	0.80	0.71	0.84	0.56	0.46	0.55	0.47
	R122A	0.43	0.71	0.49	0.86	0.47	0.94	0.31	0.70	0.47	0.77	0.58	0.58	0.45	0.61	0.50
	P124A	0.53	0.51	0.64	0.59	0.59	0.66	0.64	0.35	0.46	0.84	0.49	0.47	0.48	0.46	0.45
	Y132A	0.55	0.65	0.47	0.78	0.76	0.77	0.72	0.78	0.67	0.63	0.31	0.51	0.53	0.55	0.49
	R137A	0.68	0.45	0.58	0.49	0.34	0.72	0.75	0.80	0.69	0.68	0.77	0.54	0.65	0.72	0.64
	W139A	0.64	0.35	0.46	0.52	0.48	0.46	0.28	0.44	0.37	0.42	0.65	0.49	0.38	0.59	0.32
	Y156A	0.48	0.74	0.43	0.59	0.46	0.72	0.55	0.55	0.33	0.41	0.29	0.45	0.40	0.34	0.56
	T158A	0.88	0.79	0.52	0.91	0.83	0.42	0.82	0.82	0.51	0.78	0.50	0.62	0.67	0.74	0.66
	A159A	0.40	0.30	0.61	0.67	0.68	0.97	0.30	0.89	0.53	0.42	0.41	0.71	0.64	0.50	0.63
	S160A	0.58	0.64	0.63	0.72	0.80	0.80	0.73	0.36	0.38	0.31	0.76	0.56	0.47	0.36	0.34
	N182A	0.53	0.89	0.78	0.71	0.51	0.50	0.47	0.48	0.43	0.59	0.25	0.41	0.35	0.31	0.37
membrane-proximal	**Y202A**	0.31	0.34	0.33	0.93	0.74	0.93	0.64	0.58	0.32	0.81	0.40	0.52	0.60	0.26	0.54
	F204A	0.45	0.45	0.51	0.48	0.73	0.80	0.52	0.78	0.34	0.61	0.49	0.47	0.37	0.30	0.58

mutants (11)	**H215A**	0.70	0.69	0.76	0.97	0.74	0.96	0.54	0.80	0.68	0.73	0.52	0.53	0.49	0.43	0.51
	N225A	0.78	0.94	0.50	0.95	0.82	0.84	0.81	0.72	0.82	0.81	0.93	0.52	0.91	0.63	0.50
	E230A	0.52	0.58	0.77	0.49	0.62	0.61	0.84	0.85	0.42	0.48	0.31	0.62	0.32	0.62	0.37
	F233A	0.75	0.91	0.50	0.97	0.91	0.98	0.72	0.83	0.51	0.85	0.39	0.55	0.59	0.67	0.64
	Y242A	0.73	0.95	0.83	0.88	0.38	0.36	0.82	0.43	0.34	0.72	0.31	0.30	0.21	0.39	0.32
	K249A	0.48	0.82	0.43	0.76	0.77	0.95	0.81	0.76	0.35	0.23	0.23	0.67	0.49	0.44	0.38
	M251A	0.59	0.94	0.74	0.96	0.72	0.98	0.78	0.40	0.79	0.15	0.65	0.42	0.68	0.41	0.67
	Y253A	0.55	0.52	0.40	0.82	0.59	0.81	0.65	0.61	0.56	0.55	0.40	0.53	0.49	0.52	0.51
	T266A	0.76	0.52	0.75	0.41	0.31	0.46	0.53	0.39	0.71	0.43	0.27	0.76	0.35	0.52	0.63

Strains		**Standard error [c]**														
		Glutamic acid	**Aspartic acid**	**Tyrosine**	**Malic acid**	**Citric acid**	**Succinic acid**	**Fumaric acid**	**Oxalic acid**	**3-Hydroxypropionic acid**	**Glyceric acid**	**Dehydroasscorbic acid**	**Ribose**	**Fucose**	**Mannose**	**Ribitol**
SQR9		0.02	0.08	0.03	0.01	0.02	0.00	0.02	0.03	0.04	0.02	0.05	0.03	0.03	0.03	0.04
SQR9Δ*8mcp/mcpA*, C-*mcpA*		0.02	0.02	0.03	0.00	0.01	0.00	0.01	0.02	0.03	0.02	0.04	0.02	0.08	0.02	0.06
Chimera mutant strains (3)	***mcpA*_intracellular, M0**	0.02	0.03	0.02	0.03	0.03	0.03	0.02	0.02	0.03	0.03	0.03	0.02	0.02	0.02	0.02
	***mcpA*_distal, M1**	0.05	0.03	0.02	0.01	0.03	0.01	0.04	0.04	0.01	0.02	0.02	0.03	0.04	0.04	0.02
	***mcpA*_proximal, M2**	0.04	0.03	0.01	0.01	0.02	0.03	0.02	0.03	0.03	0.02	0.04	0.03	0.02	0.02	0.02
membrane-distal mutants (12)	**I111A**	0.03	0.03	0.03	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.04	0.01	0.03	0.05
	Y112A	0.01	0.02	0.02	0.01	0.02	0.02	0.03	0.01	0.04	0.03	0.02	0.03	0.01	0.05	0.04
	R122A	0.03	0.06	0.03	0.02	0.02	0.01	0.02	0.02	0.03	0.02	0.02	0.05	0.02	0.03	0.03
	P124A	0.03	0.01	0.02	0.02	0.01	0.03	0.02	0.03	0.03	0.02	0.01	0.03	0.02	0.02	0.03
	Y132A	0.02	0.01	0.02	0.02	0.03	0.02	0.01	0.01	0.04	0.01	0.02	0.01	0.02	0.02	0.02
	R137A	0.02	0.01	0.02	0.06	0.03	0.02	0.06	0.03	0.02	0.05	0.02	0.01	0.04	0.02	0.02
	W139A	0.02	0.03	0.03	0.01	0.04	0.02	0.03	0.03	0.02	0.04	0.02	0.02	0.03	0.03	0.03
	Y156A	0.04	0.03	0.02	0.02	0.03	0.03	0.02	0.04	0.05	0.04	0.05	0.02	0.03	0.03	0.02
	T158A	0.02	0.02	0.04	0.01	0.03	0.03	0.01	0.02	0.02	0.03	0.03	0.04	0.04	0.03	0.03
	A159A	0.03	0.02	0.03	0.03	0.02	0.02	0.05	0.02	0.04	0.07	0.02	0.02	0.02	0.04	0.03

	S160A	0.02	0.03	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.02	0.04	0.04	0.01	0.02	0.03
	N182A	0.02	0.01	0.03	0.04	0.02	0.02	0.02	0.03	0.02	0.01	0.02	0.03	0.04	0.03	0.02
membrane-proximal mutants (11)	**Y202A**	0.02	0.03	0.02	0.01	0.07	0.01	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.04
	F204A	0.02	0.03	0.02	0.03	0.02	0.02	0.03	0.02	0.03	0.05	0.02	0.04	0.02	0.03	0.02
	H215A	0.02	0.02	0.02	0.00	0.07	0.01	0.02	0.02	0.02	0.02	0.04	0.03	0.02	0.02	0.03
	N225A	0.02	0.01	0.03	0.01	0.02	0.02	0.02	0.03	0.02	0.03	0.01	0.02	0.01	0.02	0.03
	E230A	0.03	0.02	0.02	0.03	0.03	0.03	0.03	0.02	0.03	0.02	0.02	0.03	0.03	0.02	0.03
	F233A	0.02	0.01	0.02	0.01	0.01	0.00	0.03	0.01	0.03	0.00	0.05	0.03	0.03	0.03	0.04
	Y242A	0.03	0.01	0.03	0.03	0.02	0.03	0.03	0.03	0.05	0.02	0.03	0.03	0.03	0.03	0.02
	K249A	0.03	0.02	0.03	0.02	0.02	0.01	0.02	0.02	0.04	0.03	0.04	0.05	0.03	0.02	0.03
	M251A	0.03	0.01	0.03	0.00	0.03	0.00	0.02	0.03	0.03	0.03	0.03	0.02	0.05	0.03	0.02
	Y253A	0.02	0.03	0.02	0.03	0.02	0.03	0.03	0.02	0.03	0.02	0.03	0.03	0.04	0.02	0.03
	T266A	0.03	0.02	0.03	0.02	0.03	0.04	0.03	0.05	0.05	0.04	0.04	0.04	0.03	0.04	0.05

Note: [a] Compounds were used at a concentration of 1mM. [b] An I_{30} between 0.4 and 0.6 indicated no response. [c] Data for each compound are the means of thirteen biological replicates. Error bars represent the standard error of the mean.

Table S3. Experimental conditions for the initial screening of crystallisation conditions and the final crystallisation conditions.

Date (Begin)	Protein	Ligand	Kit	Protein Buffer	Result (Deadline)
07/23/2019	McpA-LBD(A36-L277)	50 mM Sodium citrate(pH 8.0)	Index/ Crystal Screen / JCSG II / JCSG IV Kits	50 mM Tris (pH 8.0), 500 mM NaCl, 1 mM TCEP, 1 mM PMSF	No crystal obtained (10/31/2019)
08/08/2019	McpA-LBD(A36-L277)	50 mM Sodium citrate(pH 8.0)	PEG/ PEGs / PACT Kits	50 mM Tris (pH 8.0), 150 mM NaCl, 1 mM TCEP, 1 mM PMSF	No crystals obtained (10/31/2019)
08/22/2019	McpA-LBD(Q38-T266)	n.a.	PEG/ PEGs / PACT Kits	50 mM Tris (pH 8.0), 150 mM NaCl, 1 mM TCEP, 1 mM PMSF	No crystal obtained (10/31/2019)
09/18/2019	McpA-LBD(A36-L277)	Apo/ 50 mM Sodium citrate (pH 8.0)	PEG-E12	50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM TCEP, 1 mM PMSF	No crystal obtained (10/31/2019)
10/29/2019	McpA-LBD(Q38-T266)-His	5mM DL-Malic acid (pH 7.6)	PEG/ PEGs / PACT/ PEGs II / JCSG I / Index / Crystal Kits	50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM TCEP, 1 mM PMSF, 5% glycerol	No crystal obtained (12/02/2019)
11/13/2019	McpA-LBD(Q38-T266)-His	5mM DL-Malic acid (pH 7.6)	JCSG II/ JCSG IV / Natrix / SaltRx Kits	50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM TCEP, 1 mM PMSF, 5% glycerol	No crystal obtained (12/02/2019)
	McpA-LBD(Q38-L224)	5mM DL-Malic acid (pH 7.6)	JCSG IV-H3	50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM TCEP, 1 mM PMSF, 5% glycerol	New crystal obtained (08/11/2020)

Note, each kit has 96 conditions, and each protein has been tried in more than 8 plates to screen the crystallization conditions.

Table S4. The summary of the data collection and refinement statistics of McpA-sLBD.

Crystal structure	**McpA-sLBD**
Data collection	
Beam line	BL17U1 (SSRF)
Detector	Eiger 16M (DECTRIS)
Temperature (K)	100
Mosaicity (°)	0.40-0.76
Wavelength (Å)	0.97918
Resolution range (Å) [a]	50.00-2.25 (2.29-2.25)
Total Number of reflections	133575
Space group	$P2_12_12$
Molecules per asymmetric unit	2
a, b, c (Å)	180.161 32.855 57.903
α, β, γ (°)	90 90 90
No. of unique reflections [a]	17195 (821)
Redundancy [a]	7.8 (5.5)
Completeness (%) [a]	99.7 (96.6)
I/sigma (I)	20.4 (3.33)
R_{merge} (%) [a]	9.1 (47.7)
CC1/2 [a]	0.996 (0.873)
Refinement	
R_{work}	0.20936
R_{free}	0.25324
Number of non-hydrogen atoms	2366
Proteins	2305
Ligands	18
Waters	43
Protein residues	187
RMS (bonds) (Å)	0.004
RMS (Angles) (°)	1.238
Ramachandran favored (%)	96.68
Allowed (%)	3.32
Outliers (%)	0.00
Average B factors	
Proteins	49.729
Ligands	56.040
Water	49.790
Starting model	6pxy
PDB ID	7w0w

[a] Numbers in parentheses represent values in the highest resolution shell.

Table S5. Summary of the prediction of key amino acids in McpA-LBD.

MCPs	Strains	PDB	Ligands	Score	Qurey cover	E-value	Identidy	Length	*in silico* analysis	Source
CtaA	*Pseudomonas fluorescens*	6pxy	L-alanine	79.0	0.44	4.0E-17	0.336	250	Eight amino acid residues: I111, R122, Y132, R137, W139, Y156, A159, N182	NCBI (PDB Database)
PctA	*Pseudomonas aeruginosa*	5ltx	L-methionine	78.2	0.44	1.0E-16	0.336	270	Seven amino acid residues: I111, R122, Y132, R137, W139, A159, N182	
PctB	*Pseudomonas aeruginosa*	5ltd	L-arginine	72.8	0.44	1.0E-14	0.291	290	Nine amino acid residues: I111, R122, P124, Y132, R137, W139, T158, A159, N182	
PctC	*Pseudomonas aeruginosa*	5ltv	γ-aminobutyric acid	68.6	0.50	3.0E-13	0.312	273	Eight amino acid residues: I111, R122, Y132, R137, W139, Y156, A159, N182	
Mcp_N	*Vibrio cholerae*	3c8c	L-alanine	59.3	0.60	4.0E-10	0.266	240	Eight amino acid residues: I111, R122, Y132, R137, W139, Y156, A159, N182	
Mlp37	*Vibrio cholerae*	5ave	serine	59.3	0.60	6.0E-10	0.266	257	Seven amino acid residues: I111, R122, Y132, R137, W139, Y156, A159	
Mlp37	*Vibrio cholerae*	6iov	arginine	58.9	0.60	6.0E-10	0.266	257	Ten amino acid residues: I111, R122, P124, Y132, R137, W139, Y156, A159, S160, N182	
Mlp24	*Vibrio cholerae*	6iop	*apo*-	52.8	0.47	8.0E-08	0.320	256	None	
CcmL	*Campylobacter jejuni* 1564	4xmr	isoleucine						Seven amino acid residues: I111, R122, Y132, R137, W139, Y156, A159	(1, 2)
PctA	*Pseudomonas aeruginosa*	5t65	isoleucine						Nine amino acid residues: I111, R122, P124, Y132, R137, W139, Y156, A159, N182	(3, 4)
TlpC	*Helicobacter pylori* 26695	5wbf	lactate						Ten amino acid residues: Y202, F204, N225, E230, F233, Y242, K249, M251, Y253, T266	(5)
PctA	*Pseudomonas aeruginosa*	5t65	acetate						Eight amino acid residues: Y202, F204, H215, N225, E230, M251, Y253, T266	(3, 4)

References for Table S5

1. H. Rahman, et al., Characterisation of a multi-ligand binding chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. PLoS Pathog 10, e1003822 (2014).
2. Y. C. Liu, M. A. Machuca, S. A. Beckham, M. J. Gunzburg, A. Roujeinikova, Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. Acta Crystallogr. Sect. D Biol. Crystallogr. 71, 2127–2136 (2015).
3. M. Rico-Jiménez, 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).
4. J. A. Gavira, et al., How bacterial chemoreceptors evolve novel ligand specificities. MBio 11, e03066-19 (2020).
5. M. A. Machuca, et al., *Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate. Sci Rep 7, 14089 (2017).

Table S6. The analysis of the significance of the differences between the chemotaxis indices of the different mutant strains as compared to the reference strain SQR9Δ*8mcp/mcpA*.

Mutant		Glutamic acid	Aspartic acid	Tyrosine	Malic acid	Citric acid	Succinic acid	Fumaric acid	Oxalic acid	3-Hydroxypropionic acid	Glyceric acid	Dehydroasscorbic acid	Ribose	Fucose	Mannose	Ribitol
membrane-distal mutants (12)	**I111A**	*N.S.*	*N.S.*	*N.S.*	***	***	***	***	***	*N.S.*	*	*N.S.*	*N.S.*	*N.S.*	*N.S.*	*
	Y112A	***	**	*	***	*N.S.*	***	**	*N.S.*	**	*N.S.*	**	*N.S.*	*	*N.S.*	*
	R122A	***	*N.S.*	*N.S.*	***	***	***	***	***	**	*N.S.*	**	*N.S.*	*	*N.S.*	*
	P124A	*	***	*N.S.*	***	***	***	***	***	***	**	***	*	*	***	**
	Y132A	*N.S.*	***	*	***	***	***	***	***	*N.S.*	***	***	*N.S.*	*N.S.*	***	*
	R137A	**	***	*N.S.*	***	***	***	**	**	*N.S.*	*N.S.*	*N.S.*	*N.S.*	*N.S.*	*	*N.S.*
	W139A	*N.S.*	***	*	***	***	***	***	***	***	***	*N.S.*	*	**	*	***
	Y156A	*	*N.S.*	**	***	***	***	***	***	***	***	***	**	**	***	*N.S.*
	T158A	***	*N.S.*	*N.S.*	***	*	***	***	**	**	*N.S.*	**	*N.S.*	*N.S.*	*	*N.S.*
	A159G	***	***	*N.S.*	***	***	*N.S.*	***	*N.S.*	*N.S.*	***	***	***	*N.S.*	**	*N.S.*
	S160A	*N.S.*	**	*N.S.*	***	***	***	***	***	***	***	*N.S.*	*N.S.*	*	***	***
	N182A	*	***	***	***	***	***	***	***	***	***	***	***	**	***	***
membrane-proximal mutants (11)	**Y202A**	***	***	***	***	*	***	***	***	***	*N.S.*	***	*N.S.*	*N.S.*	***	*N.S.*
	F204A	***	***	*N.S.*	***	***	***	***	***	***	*	***	*N.S.*	***	***	*N.S.*
	H215A	***	*N.S.*	***	***	*	***	***	***	*N.S.*	*N.S.*	**	*N.S.*	*	***	*
	N225A	***	***	*N.S.*	***	**	***	***	***	***	*N.S.*	***	*N.S.*	**	*N.S.*	*
	E230A	*	***	***	***	***	***	**	*N.S.*	***	***	***	*N.S.*	***	*N.S.*	***
	F233A	***	***	*N.S.*	*N.S.*	*N.S.*	*N.S.*	***	***	*	***	**	*N.S.*	*N.S.*	*N.S.*	*N.S.*
	Y242A	**	***	***	**	***	***	***	***	***	*N.S.*	***	***	***	***	***
	K249A	***	*	**	***	***	*	***	***	***	***	***	*N.S.*	*	***	***
	M251A	*N.S.*	***	***	***	***	*N.S.*	***	***	***	***	*N.S.*	***	*N.S.*	***	*N.S.*
	Y253A	*N.S.*	***	***	***	***	***	***	***	*N.S.*	***	***	*N.S.*	*N.S.*	***	*
	T266A	***	***	***	***	***	***	***	***	*N.S.*	***	***	***	**	*	*N.S.*
chimera mutants	**McpA-intracellular**	*	***	*	***	***	***	***	***	***	***	**	***	*	***	**
	McpA-distal	*N.S.*	***	*N.S.*	**	***	***	**	***	***	*N.S.*	*N.S.*	*N.S.*	*	*N.S.*	*

	McpA-proximal	**	*N.S.*	***	***	*	***	***	***	*N.S.*	***	***	***	*	**	*N.S.*

Note: The Duncan's multiple rang tests ($p < 0.05$) of the SPSS version 22.0 (IBM, Chicago, IL, version 22.0) was used for statistical analysis. *N.S.* refers to no significant; * $0.01 < p < 0.05$; ** $0.001 < p < 0.01$; *** $p < 0.001$.

Table S7. The analysis of key amino acid residues in sensing the different chemoeffectors. Shown are the number of compounds of a given class that present a given chemotaxis phenotype.

Mutant		Amino acids (3)				Organic acids (8)				Sugar and sugar alcohols (4)				All the compounds (15)				
		No chemotaxis	Repellent	Reduced	Enhanced	No chemotaxis	Repellent	Reduced	Enhanced	No chemotaxis	Repellent	Reduced	Enhanced	No chemotaxis	Repellent	Reduced	Enhanced	**Total**
membrane-distal mutants (12)	I111A	0	0	0	0	0	0	6	0	1	0	0	0	1	0	6	0	**7**
	Y112A	1	0	0	2	0	0	3	2	2	0	0	0	3	0	3	4	**10**
	R122A	1	0	0	0	3	1	3	0	2	0	0	0	6	1	3	0	**10**
	P124A	2	0	0	0	4	1	2	1	4	0	0	0	10	1	2	1	**14**
	Y132A	1	0	1	0	0	1	6	0	2	0	0	0	3	1	7	0	**11**
	R137A	1	0	0	1	1	1	3	0	0	0	0	1	2	1	3	2	**8**
	W139A	1	1	0	0	5	2	0	0	2	2	0	0	8	5	0	0	**13**
	Y156A	2	0	0	0	5	2	1	0	2	1	0	0	9	3	1	0	**13**
	T158A	0	0	0	1	3	0	4	0	0	0	0	1	3	0	4	2	**9**
	A159A	1	1	0	0	2	1	2	0	1	0	0	1	4	2	2	1	**9**
	S160A	0	0	1	0	0	3	4	0	1	2	0	0	1	5	5	0	**11**
	N182A	1	0	0	2	6	1	1	0	1	3	0	0	8	4	1	2	**15**
membrane-proximal mutants (11)	Y202A	0	3	0	0	2	1	4	0	0	1	0	0	2	5	4	0	**11**
	F204A	2	0	0	0	3	1	4	0	0	2	0	0	5	3	4	0	**12**
	H215A	0	0	0	2	2	0	4	0	3	0	0	0	5	0	4	2	**11**
	N225A	0	0	0	2	0	0	5	2	1	0	0	1	1	0	5	5	**11**
	E230A	2	0	0	1	3	1	3	0	0	2	0	0	5	3	3	1	**12**
	F233A	0	0	0	2	1	1	2	1	0	0	0	0	1	1	2	3	**7**
	Y242A	0	0	0	3	1	4	2	0	0	4	0	0	1	8	2	3	**14**
	K249A	2	0	0	1	0	3	5	0	2	1	0	0	4	4	5	1	**14**
	M251A	0	0	0	2	1	1	3	1	2	0	0	0	3	1	3	3	**10**
	Y253A	1	1	0	0	2	1	4	0	2	0	0	0	5	2	4	0	**11**
	T266A	1	0	0	2	4	3	0	0	1	1	0	1	6	4	0	3	**13**

Table S8. The difference of chemotaxis indices of the individual mutants with respect to the reference strain SQR9Δ*8mcp/mcpA* for 15 different chemoeffectors.

Mutant		Amino acids (3)			Organic acids (8)								Sugar and sugar alcohols (4)			
		Glutamic acid	Aspartic acid	Tyrosine	Malic acid	Citric acid	Succinic acid	Fumaric acid	Oxalic acid	3-Hydroxypropionic acid	Glyceric acid	Dehydroascorbic acid	Ribose	Fucose	Mannose	Ribitol
membrane-distal mutants (12)	I111A	0.020	-0.053	-0.057	-0.216	-0.263	-0.187	-0.228	-0.230	-0.000	-0.075	0.072	0.047	-0.051	-0.032	-0.163
	Y112A	0.303	0.099	-0.078	-0.092	-0.025	-0.179	-0.120	-0.012	0.172	-0.039	0.129	-0.005	-0.209	-0.103	-0.178
	R122A	-0.167	-0.039	-0.074	-0.131	-0.449	-0.046	-0.627	-0.205	-0.161	0.024	-0.123	0.012	-0.223	-0.041	-0.153
	P124A	-0.070	-0.239	0.079	-0.398	-0.329	-0.323	-0.301	-0.551	-0.165	0.092	-0.214	-0.095	-0.192	-0.195	-0.199
	Y132A	-0.045	-0.100	-0.095	-0.204	-0.158	-0.221	-0.222	-0.121	0.043	-0.119	-0.394	-0.053	-0.143	-0.103	-0.158
	R137A	0.084	-0.296	0.020	-0.501	-0.571	-0.271	-0.186	-0.100	0.059	-0.074	0.067	-0.021	-0.019	0.072	-0.004
	W139A	0.045	-0.396	-0.103	-0.470	-0.434	-0.529	-0.664	-0.467	-0.258	-0.330	-0.055	-0.075	-0.297	-0.065	-0.324
	Y156A	-0.118	-0.009	-0.129	-0.399	-0.460	-0.272	-0.392	-0.354	-0.295	-0.336	-0.420	-0.121	-0.270	-0.311	-0.091
	T158A	0.279	0.046	-0.045	-0.082	-0.084	-0.568	-0.119	-0.084	-0.117	0.029	-0.208	0.052	-0.001	0.092	0.017
	A159A	-0.193	-0.446	0.051	-0.317	-0.232	-0.019	-0.639	-0.016	-0.102	-0.325	-0.302	0.142	-0.030	-0.151	-0.017
	S160A	-0.019	-0.107	0.071	-0.268	-0.116	-0.188	-0.206	-0.543	-0.251	-0.445	0.053	-0.001	-0.200	-0.296	-0.307
	N182A	-0.066	0.145	0.213	-0.273	-0.402	-0.484	-0.474	-0.427	-0.194	-0.160	-0.458	-0.158	-0.324	-0.342	-0.281
membrane-proximal mutants (11)	Y202A	-0.288	-0.408	-0.231	-0.061	-0.173	-0.062	-0.301	-0.330	-0.310	0.057	-0.307	-0.049	-0.069	-0.394	-0.103
	F204A	-0.149	-0.291	-0.054	-0.503	-0.191	-0.184	-0.422	-0.128	-0.291	-0.141	-0.219	-0.096	-0.304	-0.347	-0.065
	H215A	0.108	-0.056	0.198	-0.018	-0.173	-0.030	-0.404	-0.105	0.048	-0.025	-0.192	-0.034	-0.182	-0.223	-0.137
	N225A	0.187	0.191	-0.060	-0.039	-0.097	-0.148	-0.129	-0.185	0.190	0.055	0.219	-0.047	0.233	-0.019	-0.145
	E230A	-0.074	-0.164	0.209	-0.497	-0.293	-0.377	-0.105	-0.054	-0.212	-0.273	-0.401	0.051	-0.351	-0.034	-0.278
	F233A	0.152	0.168	-0.061	-0.015	-0.002	-0.004	-0.219	-0.077	-0.120	0.099	-0.321	-0.013	-0.080	0.013	-0.006
	Y242A	0.136	0.204	0.269	-0.105	-0.532	-0.631	-0.119	-0.475	-0.284	-0.033	-0.398	-0.266	-0.467	-0.264	-0.330
	K249A	-0.121	0.078	-0.132	-0.225	-0.144	-0.034	-0.126	-0.141	-0.281	-0.521	-0.477	0.104	-0.183	-0.216	-0.264
	M251A	-0.009	0.191	0.173	-0.026	-0.195	-0.005	-0.164	-0.501	0.161	-0.598	-0.056	-0.149	0.012	-0.241	0.025
	Y253A	-0.047	-0.229	-0.168	-0.169	-0.328	-0.181	-0.286	-0.292	-0.064	-0.195	-0.310	-0.041	-0.178	-0.128	-0.139
	T266A	0.165	-0.225	0.188	-0.582	-0.610	-0.527	-0.414	-0.518	0.083	-0.316	-0.440	0.190	-0.317	-0.128	-0.022

Note: black and red represent an increase and decrease in the chemotaxis index, respectively.

Table S9. Summary of the ligands recognized by typical dCache containing chemoreceptors in *B. velezensis* SQR9 and other strains.

Genus	Species	MCP	Identity [a]	Ligands	References
Bacillus	*B. velezensis* SQR9	McpA	100%	**aspartic acid**, glutamic acid, **isoleucine**, **lysine**, **tyrosine**; phthalic acid, citric acid, oxalic acid, **malic acid**, **succinic acid**, **fumaric acid**, **adipic acid**, dehydroascorbic acid, **glyceric acid**, **3-hydroxypropionic acid**; ribitol, mannose, **ribose**, **fucose**, **hydroxycarbamate**, **sodium decanoate**, **serine**, **gluconic acid**, **fructose, galactose**	(1–4)
		TlpA	47.3%	dithiothreitol, maltose, **hydroxycarbamate, sodium decanoate, gluconic acid**	
		TlpB	43.5%	phenylalanine, dulcitol, pentadecanoic acid, inosine, maltose, salicylic acid, threonine, **fructose, malic acid, succinic acid, fumaric acid, ribose, fucose, sodium decanoate, gluconic acid**	
		McpB	40.3%	glycine, tryptophan, asparagine, glutamine, salicylic acid, serine, cystine, methionine, **sodium decanoate, adipic acid, ribose, glyceric acid, 3-hydroxypropionic acid, gluconic acid, fructose**	
		McpC	21.4%	valine, alanine, proline, leucine, histidine, threonine, cystine, methionine, maltose, **serine, gluconic acid, succinic acid**	
	B. subtilis OI1085	McpA	67.0%	glucose, alpha-D-ethylglucoside, acidic environments	(4–8)
		TlpB	38.8%	alkaline environments	
		TlpA	38.8%	acidic environments	
		McpB	38.0%	asparagine, **aspartic acid**, glutamine, histidine; methanol, alkaline environments	
		McpC	18.2%	15 amino acids (**serine, tyrosine, isoleucine, lysine**)	
Pseudomonas	*P. aeruginosa* PAO1	PctA	19.5%	17 amino acids (**serine, tyrosine, isoleucine, lysine**), autoinducer-2, histamine,	(4, 9–13)

				glucose (gltB)	
		PctB	17.0%	arginine, alanine, methionine, glutamine, **lysine**	
		PctC	18.2%	proline, GABA, histidine	
		PctD	12.69%	acetylcholine, choline	
		TlpQ	9.42%	putrescine, cadaverine, spermidine, agmatine, ethylenediamine, histamine, autoinducer-2	
	P. fluorescens Pf0-1	CtaA	20.2%	16 amino acids (**serine, tyrosine, isoleucine, lysine**)	(4, 14)
	P. putida KT2440	McpA	21.6%	12 amino acids (**serine, tyrosine, isoleucine**)	(4, 10, 15–18)
		McpG	20.39%	GABA	
		McpH	16.3%	uric acid, adenine, guanine, hypoxanthine, purine, xanthine	
		McpU	11.88%	putrescine, cadaverine, spermidine, agmatine, ethylenediamine, histamine	
Helicobacter	*H. pylori* SS1	TlpA	13.0%	arginine, cysteine, glucosamine, **fumarate**	(19, 20)
	H. pylori 26695	TlpC	14.2%	lactate	
Sinorhizobium	*S. meliloti* MVII-1	McpU	19.4%	proline	(21, 22)
		McpX	16.8%	quaternary ammonium compounds	
Vibrio	*V. cholerae* O395N1	Mlp37	18.7%	taurine, alanine, arginine, **serine**	(23)
Campylobacter	*C. jejuni* 520	Tlp11	14.1%	**galactose**	(24)

Note: [a] The sequence identity refers to the similarities between MCP-LBD and the targeted McpA-LBD of SQR9. The LBDs were predicted by the SMART online tool (http://smart.embl-heidelberg.de/).

Bold red font indicates that these compounds were also sensed by the McpA in *B. velezensis* SQR9.

References for Table S9

1. H. Feng, et al., Identification of chemotaxis compounds in root exudates and their sensing chemoreceptors in plant-growth-promoting rhizobacteria *Bacillus amyloliquefaciens* SQR9. Mol. Plant-Microbe Interact. 31, 995-1005 (2018).
2. H. Feng, et al., Recognition of dominant attractants by key chemoreceptors mediates recruitment of plant growth-promoting rhizobacteria. Env. Microbiol 21, 402-415 (2019).
3. Y. Liu, et al., Induced root-secreted D-galactose functions as a chemoattractant and enhances the biofilm formation of *Bacillus velezensis* SQR9 in an McpA-dependent manner. Appl. Microbiol. Biotechnol. 104, 785-797 (2020).
4. H. Feng, et al., Chemotaxis of beneficial rhizobacteria to root exudates: The first step towards root-microbe rhizosphere interactions. Int. J. Mol. Sci. 22, 6655 (2021).
5. D. W. Hanlon, G. W. Ordal, Cloning and characterization of genes encoding methyl-accepting chemotaxis proteins in *Bacillus subtilis*. J Biol Chem 269, 14038-14046 (1994).
6. G. D. Glekas, et al., The *Bacillus subtilis* chemoreceptor McpC senses multiple ligands using two discrete mechanisms. J. Biol. Chem. 287, 39412-39418 (2012).
7. P. Tohidifar, M. J. Plutz, G. W. Ordal, C. V. Rao, The mechanism of bidirectional pH taxis in *Bacillus subtilis*. J. Bacteriol. 202, e00491-19 (2020).
8. P. Tohidifar, et al., The unconventional cytoplasmic sensing mechanism for ethanol chemotaxis in *Bacillus subtilis*. MBio 11, e02177-20 (2020).
9. M. Rico-Jimenez, 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).
10. A. Corral-Lugo, et al., High-affinity chemotaxis to histamine mediated by the TlpQ chemoreceptor of the human pathogen *Pseudomonas aeruginosa*. MBio 9, e01894-18 (2018).
11. L. Zhang, et al., Sensing of autoinducer-2 by functionally distinct receptors in prokaryotes. Nat. Commun. 11, 5371 (2020).
12. C. Xu, Q. Cao, L. Lan, Glucose-binding of periplasmic protein GltB activates GtrS-GltR two-component system in *Pseudomonas aeruginosa*. Microorganisms 9, 447 (2021).
13. M. A. Matilla, et al., Chemotaxis of the human pathogen *Pseudomonas aeruginosa* to the neurotransmitter acetylcholine. MBio 13, e0345821 (2022).
14. A. I. M. S. Ud-Din, M. F. Khan, A. Roujeinikova, Broad specificity of amino acid chemoreceptor CtaA of *Pseudomonas fluorescens* is afforded by plasticity of its amphipathic ligand-binding pocket. Mol. Plant-Microbe Interact. 33, 612-623 (2020).
15. J. A. Reyes-Darias, et al., Specific gamma-aminobutyrate chemotaxis in pseudomonads with different lifestyle. Mol Microbiol 97, 488-501 (2015).
16. M. Fernández, B. Morel, A. Corral-Lugo, T. Krell, Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. Mol. Microbiol. 99, 34-42 (2016).
17. A. Corral-Lugo, et al., Assessment of the contribution of chemoreceptor-based signalling to biofilm formation. Environ. Microbiol. 18, 3355-3372 (2016).

18. J. A. Gavira, et al., Structural basis for polyamine binding at the dCACHE domain of the McpU chemoreceptor from *Pseudomonas putida*. J. Mol. Biol. 430, 1950-1963 (2018).
19. M. A. Machuca, et al., *Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate. Sci Rep 7, 14089 (2017).
20. K. S. Johnson, et al., The dCache chemoreceptor TlpA of *Helicobacter pylori* binds multiple attractant and antagonistic ligands via distinct sites. MBio 12, e01819-21 (2021).
21. B. A. Webb, S. Hildreth, R. F. Helm, B. E. Scharf, *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. Appl. Environ. Microbiol. 80, 3404-3415 (2014).
22. B. A. Webb, et al., *Sinorhizobium meliloti* chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX. Mol Microbiol 103, 333-346 (2017).
23. S. Nishiyama, et al., Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. Sci. Rep. 6, 20866 (2016).
24. C. J. Day, et al., A direct-sensing galactose chemoreceptor recently evolved in invasive strains of *Campylobacter jejuni*. Nat. Commun. 7, 13206 (2016).

Table S10. Primers used in the study.

Primer name	Sequence (5' to 3')	Purpose
mcpA_intracelluar-amyE-UF	TCTCGATAATATGGTAGGC	Complementation of the *mcpA*_intracelluar gene
mcpA_intracelluar-amyE-UR	TTTGCTGTTTTTTTCAACTGCTATATCTCCTCTTGC	
mcpA_intracelluar-F	GCAAGAGGAGATATAGCAGTTGAAAAAAACAGCAAA	
mcpA_intracelluar-R	AGCAGGGACAGATGAAGGATCGGTCTTGCCGCATCCTTGATTT	
mcpA_intracelluar-amyE-DF	AAATCAAGGATGCGGCAAGACCGATCCTTCATCTGTCCCTGCT	
mcpA_intracelluar-amyE-DR	GAGAAGGCGTCGTAAACT	
mcpA_intracelluar-VF	ATGTGATGATGGCCCTAT	
mcpA_intracelluar-VR	GGAAGATGCGACGTTATT	
mcpA_distal-amyE-UF	TCTCGATAATATGGTAGGC	Complementation of the *mcpA*_distal gene
mcpA_distal-amyE-UR	CATCGTGATGAGATCGTCGATTTTCATATTGATGGC	
mcpA_distal-F	GCCATCAATATGAAAATCGACGATCTCATCACGATG	
mcpA_distal-R	CAGGGACAGATGAAGGATCGGTCTTGCCGCATCCT	
mcpA_distal-amyE-DF	AGGATGCGGCAAGACCGATCCTTCATCTGTCCCTG	
mcpA_distal-amyE-DR	GAGAAGGCGTCGTAAACT	
mcpA_distal-VF	AGCGACTGGCTCACAAAG	
mcpA_distal-VR	AAGCCGTCAGCTCCTCTG	
mcpA_proximal-amyE-UF	TCTCGATAATATGGTAGGC	Complementation of the *mcpA*_proximal gene
mcpA_proximal-amyE-UR	TTTGCTGTTTTTTTCAACTGCTATATCTCCTCTTGC	
mcpA_proximal-F	GCAAGAGGAGATATAGCAGTTGAAAAAAACAGCAAA	
mcpA_proximal-R	CGTTTTCAGAAGCTCATCGATTTTTATATCAATGGC	
mcpA_proximal-amyE-DF	GCCATTGATATAAAAATCGATGAGCTTCTGAAAACG	
mcpA_proximal-amyE-DR	GAGAAGGCGTCGTAAACT	
mcpA_proximal-VF	AAGTAAGGGTTTGATGTGAT	
mcpA_proximal-VR	AAGATGCGACGTTATTCA	
Mut-Com-VF	CTGGATGGCATGGATGA	Verification of all the complementation of the *mcpA* with a single point mutant
Mut-Com-VR	GGCAGATGATGGGAAGC	
I111A-UF	GAGAAGGCGTCGTAAACTGTG	Construction of the *mcpA* gene with a point mutant
I111A-UR	ACGTGGAATCCGCCTATACGAGTGACACGGAAGGACATTTCACAC	

I111A-DF	GTGTGAAATGTCCTTCCGTGTCACTCGTATAGGCGGATTCCACGT	
I111A-DR	TCTCGATAATATGGTAGGCCGC	
Y112A-UF	GAGAAGGCGTCGTAAACTGTG	
Y112A-UR	GAAGATCAATAAAGACGTGGAATCCATCGCCACGAGTGACACGGAAGGACA	
Y112A-DF	TGTCCTTCCGTGTCACTCGTGGCGATGGATTCCACGTCTTTATTGATCTTC	
Y112A-DR	TCTCGATAATATGGTAGGCCGC	
R122A-UF	GAGAAGGCGTCGTAAACTGTG	
R122A-UR	CGGAAGGACATTTCACAGCCTATCCGGATCTGCAGATGCCTA	
R122A-DF	TAGGCATCTGCAGATCCGGATAGGCTGTGAAATGTCCTTCCG	
R122A-DR	TCTCGATAATATGGTAGGCCGC	
P124A-UF	GAGAAGGCGTCGTAAACTGTG	
P124A-UR	ACATTTCACACGGTATGCCGATCTGCAGATGCCTAAAGGATA	
P124A-DF	TATCCTTTAGGCATCTGCAGATCGGCATACCGTGTGAAATGT	
P124A-DR	TCTCGATAATATGGTAGGCCGC	
Y132A-UF	GAGAAGGCGTCGTAAACTGTG	
Y132A-UR	CTGCAGATGCCTAAAGGAGCCAATCCGATTGAAAGAGACTGGTATA	
Y132A-DF	TATACCAGTCTCTTTCAATCGGATTGGCTCCTTTAGGCATCTGCAG	
Y132A-DR	TCTCGATAATATGGTAGGCCGC	
R137A-UF	GAGAAGGCGTCGTAAACTGTG	
R137A-UR	AATCCGATTGAAGCCGACTGGTATAAAAAGGCCGTGGAA	
R137A-DF	TTCCACGGCCTTTTTATACCAGTCGGCTTCAATCGGATT	
R137A-DR	TCTCGATAATATGGTAGGCCGC	
W139A-UF	GAGAAGGCGTCGTAAACTGTG	
W139A-UR	TATAATCCGATTGAAAGAGACGCCTATAAAAAGGCCGTGGAAAAT	
W139A-DF	ATTTTCCACGGCCTTTTTATAGGCGTCTCTTTCAATCGGATTATA	
W139A-DR	TCTCGATAATATGGTAGGCCGC	
Y156A-UF	GAGAAGGCGTCGTAAACTGTG	
Y156A-UR	TGTAACAGATCCTGCCCGCACGGCTTCCACCAATACGAT	
Y156A-DF	ATCGTATTGGTGGAAGCCGTGCGGGCAGGATCTGTTACA	
Y156A-DR	TCTCGATAATATGGTAGGCCGC	
T158A-UF	GAGAAGGCGTCGTAAACTGTG	
T158A-UR	TCCTTACCGCGCCGCTTCCACCAATACGATGGTCGTGAC	

T158A-DF	GTCACGACCATCGTATTGGTGGAAGCGGCGCGGTAAGGA	
T158A-DR	TCTCGATAATATGGTAGGCCGC	
A159G-UF	GAGAAGGCGTCGTAAACTGTG	
A159G-UR	TACCGCACGGGATCCACCAATACGATGGTCGTGACAGT	
A159G-DF	ACTGTCACGACCATCGTATTGGTGGATCCCGTGCGGTA	
A159G-DR	TCTCGATAATATGGTAGGCCGC	
S160A-UF	GAGAAGGCGTCGTAAACTGTG	
S160A-UR	CGCACGGCTGCCACCAATACGATGGTCGTGACAGTC	
S160A-DF	GACTGTCACGACCATCGTATTGGTGGCAGCCGTGCG	
S160A-DR	TCTCGATAATATGGTAGGCCGC	
N182A-UF	GAGAAGGCGTCGTAAACTGTG	
N182A-UR	TCGTCGCCATCGCCATGAAAATCGATGAGCTTCTGAAAAC	
N182A-DF	GTTTTCAGAAGCTCATCGATTTTCATGGCGATGGCGACGA	
N182A-DR	TCTCGATAATATGGTAGGCCGC	
Y202A-UF	GAGAAGGCGTCGTAAACTGTG	
Y202A-UR	ATCGGAAAATCCGGCGCCGCTTTTATCCTGACGAAAGAC	
Y202A-DF	GTCTTTCGTCAGGATAAAAGCGGCGCCGGATTTTCCGAT	
Y202A-DR	TCTCGATAATATGGTAGGCCGC	
F204A-UF	GAGAAGGCGTCGTAAACTGTG	
F204A-UR	AAATCCGGCTACGCTGCCATCCTGACGAAAGACAAA	
F204A-DF	TTTGTCTTTCGTCAGGATGGCAGCGTAGCCGGATTT	
F204A-DR	TCTCGATAATATGGTAGGCCGC	
H215A-UF	GAGAAGGCGTCGTAAACTGTG	
H215A-UR	AAGTGGTCGCCGCCCCGAACCGCACGTCAGGCACAG	
H215A-DF	CTGTGCCTGACGTGCGGTTCGGGGCGGCGACCACTT	
H215A-DR	TCTCGATAATATGGTAGGCCGC	
N225A-UF	GAGAAGGCGTCGTAAACTGTG	
N225A-UR	GGCACAGTATTAGCCGGGGAGTGGGCTGAACAGCTGTTCACC	
N225A-DF	GGTGAACAGCTGTTCAGCCCACTCCCCGGCTAATACTGTGCC	
N225A-DR	TCTCGATAATATGGTAGGCCGC	
E230A-UF	GAGAAGGCGTCGTAAACTGTG	
E230A-UR	GGGGAGTGGGCTGCCCAGCTGTTCACCAGCAAAAAAGGT	

E230A-DF	ACCTTTTTTGCTGGTGAACAGCTGGGCAGCCCACTCCCC	
E230A-DR	TCTCGATAATATGGTAGGCCGC	
F233A-UF	GAGAAGGCGTCGTAAACTGTG	
F233A-UR	CTGAACAGCTGGCCACCAGCAAAAAAGGTGATTTTCAATACAAATACGAC	
F233A-DF	GTCGTATTTGTATTGAAAATCACCTTTTTTGCTGGTGGCCAGCTGTTCAG	
F233A-DR	TCTCGATAATATGGTAGGCCGC	
Y242A-UF	GAGAAGGCGTCGTAAACTGTG	
Y242A-UR	AAGGTGATTTTCAAGCCAAATACGACGGCCAGGAAAAGAAAATGGC	
Y242A-DF	GCCATTTTCTTTTCCTGGCCGTCGTATTTGGCTTGAAAATCACCTT	
Y242A-DR	TCTCGATAATATGGTAGGCCGC	
K249A-UF	GAGAAGGCGTCGTAAACTGTG	
K249A-UR	GACGGCCAGGAAGCCAAAATGGCTTATGATACAAACGGGCTGAC	
K249A-DF	GTCAGCCCGTTTGTATCATAAGCCATTTTGGCTTCCTGGCCGTC	
K249A-DR	TCTCGATAATATGGTAGGCCGC	
M251A-UF	GAGAAGGCGTCGTAAACTGTG	
M251A-UR	GGCCAGGAAAAGAAAGCCGCTTATGATACAAACGGGCTGAC	
M251A-DF	GTCAGCCCGTTTGTATCATAAGCGGCTTTCTTTTCCTGGCC	
M251A-DR	TCTCGATAATATGGTAGGCCGC	
Y253A-UF	GAGAAGGCGTCGTAAACTGTG	
Y253A-UR	GGCCAGGAAAAGAAAATGGCTGCCGATACAAACGGGCTGACAGGCTGGAAGATT	
Y253A-DF	AATCTTCCAGCCTGTCAGCCCGTTTGTATCGGCAGCCATTTTCTTTTCCTGGCC	
Y253A-DR	TCTCGATAATATGGTAGGCCGC	
T266A-UF	GAGAAGGCGTCGTAAACTGTG	
T266A-UR	CTGGAAGATTACCGGAGCCATGTATCTGGACGAAATTCATGAAGCGG	
T266A-DF	CCGCTTCATGAATTTCGTCCAGATACATGGCTCCGGTAATCTTCCAG	
T266A-DR	TCTCGATAATATGGTAGGCCGC	

Legend for Dataset S1

Dataset S1. Bioinformatics data. **Sheet 1**. The presence of McpA-LBD homologs in different bacterial strains. **Sheet 2**. The abundance of McpA-LBD homologs in different species.