

Supplementary Information for
“Discovery of Novel Chemoeffectors and Rational Design of
***Escherichia coli* Chemoreceptor Specificity”**

Shuangyu Bi^{a,b}, Daqi Yu^{a,b}, Guangwei Si^{b,c}, Chunxiong Luo^{b,c}, Tongqing Li^{a,b},
Qi Ouyang^{b,c,d}, Vladimir Jakovljevic^e, Victor Sourjik^{e,f}, Yuhai Tu^{b,g,1},
and Luhua Lai^{a,b,d,1}

^a BNLM, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.

^b Center for Quantitative Biology, AAIS, Peking University, Beijing 100871, China.

^c College of Physics, Peking University, Beijing 100871, China.

^d Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

^e Zentrum für Molekulare Biologie der Universität Heidelberg, DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany.

^f Max Planck Institute for Terrestrial Microbiology, D-35043 Marburg, Germany.

^g IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA.

¹ Corresponding authors:

Luhua Lai

College of Chemistry and Molecular Engineering, Peking University
Beijing 100871, China.

E-mail: lhlai@pku.edu.cn; Tel: 86-10-62757486

Yuhai Tu

IBM T. J. Watson Research Center, Yorktown Heights
New York 10598, USA

E-mail: yuhai@us.ibm.com; Tel: (914)945-2762

SI Results

Binding affinity measurements by ITC. ITC was performed at 25 °C on a MicroCal ITC200 calorimeter (GE Healthcare) to measure the binding affinities of compounds with the purified Tar periplasmic domain. Titrations were carried out in a buffer of 200 mM phosphate buffered saline, 200 mM NaCl. The results of ITC are shown in Table S1 and Fig. S1. As most of the compounds bind weakly, we followed the guidelines for measuring low affinity ligand binding using ITC (1-2, www.gelifesciences.com/microcal). The highest possible concentrations of the protein and compounds permitted by solubility were used while keeping a constant pH. Eight of the eleven titrations reached over 80% receptor saturation and three of them were around 70%. The resulting c values ($c = nK_d[M_0]$, M_0 is the concentration of proteins in the cell, n is the number of sites) were between 0.3 and 0.005, which were above the recommended lowest c value in ITC studies (1).

Novel chemoeffectors identified by microfluidic experiments. We discovered six attractants using microfluidic experiments. Purity analysis eliminated possible contamination of Asp in the compound samples (Fig. S2). We measured the responses of *E. coli* RP437 cells to different source concentrations of novel attractants (Fig. S3A-F). CHDCA and PA, both of which bind with Tar, did not attract cells even at high source concentration of 0.1 M (Fig. S3G-H).

FRET measurement of intracellular response to novel chemoeffectors. As

shown in Fig. S4A-G, the Tar-only strain that expresses the wild-type Tar receptor, and the CheY-YFP/CheZ-CFP FRET pair were stimulated by stepwise addition or removal of attractants at indicated concentration. Upon addition of attractants, the FRET signal (the ratio of YFP/CFP) decreased, reflecting lowered kinase activity. The Tsr-only strain, which expresses the chimera receptor of Tar and Tsr had no FRET response when stimulated by the novel attractants (Fig. S4H). PA cannot induce the change of FRET signals, as shown in Fig. S4I. The repellents, such as nickel ion, have opposite effects to attractants (Fig. S4I).

The futile binders act as antagonists that compete with attractants for binding. To make sure the antagonist does not affect the cell's swimming speed, the mean speed and mean angular speed of UU1624 swimming in the blank buffer and ambient 1 mM CHDCA were measured following previous study (3). Cells were tracked by Image J (National Institutes of Health). Data were analyzed according to previous method (3). We found that the mean speed and mean angular speed of UU1624 cells swimming in the blank buffer and in the ambient 1 mM CHDCA were almost the same during the experimental period, indicating that 1 mM CHDCA have little effect on the vitality and motility of cells (Table S2). We measured the influence of CHDCA to the intracellular response to the steps of AMA using FRET. We observed that 1 mM CHDCA could influence the FRET response of Tar-only strain to AMA. The difference of YFP/CFP change could be observed when adding AMA with 1 mM CHDCA together, as shown in Fig. S4J.

Converting an antagonist to an attractant. We measured the responses of *E. coli* cells to multiple concentration gradients of *cis*-(2*R*, 3*S*)-2,3-piperidine dicarboxylic acid (*cis*-PDA) using the microfluidic device (Fig.1A) as well as the FRET measurement. *Cis*-PDA can attract cells possessing functional Tar (Fig. S5A-E). We also measured the responses of different *E. coli* strains to multiple concentration gradients of L-malic acid (LMA) using the microfluidic device (Fig.1A). LMA can attract cells possessing functional Tar (Fig. S5F-H).

Rational design of Tar for novel chemotaxis specificity. We redesigned chemoreceptor Tar to recognize L-arginine, a basic amino acid that cannot be sensed by the wild-type Tar (Fig. S6A). The Tar mutant R69'ER73'E can sense L-arginine as an attractant. Besides L-arginine, R69'ER73'E showed weak attractant response to L-aspartate, weaker than its response to arginine (Fig. S6B). We have also verified that the mutant receptor R69'E, R73'E and R69'ER73'E have similar expression level with the wild-type Tar by using Western Blot (Fig. S6C and D).

SI Materials and Methods

Strains, plasmids, and materials. Information regarding the genotypes, phenotypes, and sources of the bacterial strains and plasmids used in this study are listed in Table S3. Guanidinosuccinic acid, (\pm)- α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid, formimino-L-aspartate, and N-formyl-L-aspartate were purchased from Sigma Aldrich. N-methyl-L-aspartate was purchased from Acros Organics. *cis* 1,2-cyclohexanedicarboxylic acid and phthalic acid were purchased

from Alfa Aesar. (2-imino-4-oxo-thiazolidin-5-yl)-acetic acid was purchased from Matrix Scientific. *cis*-(2*R*,3*S*)-2,3-piperidine dicarboxylic acid was purchased from Beijing Repharma Co., Ltd. L-malic acid was purchased from J&K Scientific Ltd.

Virtual screening for novel chemoeffectors. As far as we know, there are at least three available crystal structures (4-6) for the Tar periplasmic domain with ligands in *Salmonella* and *E. coli*. Previous studies (6) showed little differences among these structure candidates in the binding interactions with aspartate. Currently, no crystal structure is available for the *E. coli* periplasmic domain of Tar with the ligand Asp bound. The available apo structure (7) or the pseudoligand-bound structure (8) does not reflect the specific Asp binding-induced conformational changes. The sequences between the periplasmic domain of Tar in *Salmonella* and *E. coli* share 66% identity with no gaps. We chose the newly published Tar receptor structure from *Salmonella* (PDB code: 1VLT) (6) as the template to build the *E. coli* receptor structure for the virtual screening. The sequence of the Tar periplasmic domain in *Salmonella* was mutated to that in *E. coli* using Scap (9), a program for side chain conformation prediction and residue mutation. Ninety among 284 residues were mutated in total. The mutated residues were all farther than 5 Å away from the binding pocket. The mutated structure was then optimized in CHARMM c33b1 (10). The ligands and water molecules were removed in the next steps. We modeled the structure of the *E. coli* Tar periplasmic domain based on the crystal structure of *Salmonella* Tar. The AutoDock program (version 4.0.1) was used for the virtual screening by docking (11). Molecules with molecular weight < 300 Da in the MDL

ACD were selected for the docking study (149,063 molecules). This molecular weight limit was set based on the size of the aspartate binding pocket. The top 10,000 compounds with the lowest estimated binding free energies lower than $-5.5 \text{ kcal mol}^{-1}$ were selected. Eighty compounds were selected manually and purchased for experimental studies.

Clone and mutagenesis of the periplasmic domain of Tar. The plasmid pMDL101 was constructed to express the *E. coli* Tar periplasmic domain. The coding sequence of residues 32-188 was amplified using PCR reaction from the pLC113, a plasmid encodes wild-type full length *tar*. Two oligonucleotides were used for the PCR reaction, introducing restriction sites *Nde*I at the 5'-end, *Bam*HI and a stop codon TGA at the 3'-end. The amplified fragment was digested and ligated with pET-28a (His-Tag containing expression vector; Novagen) to create plasmid pMDL101. Mutants of the periplasmic domain of Tar, R64A, R69'D, and R73'A, were generated by Muta-direct™ site-directed mutagenesis kit (SBS Genetech). The plasmid pMDL101 was the template for the mutagenesis. All mutants were verified by DNA sequencing.

Expression and purification of wild-type and mutant Tar periplasmic domain. The plasmid pMDL101 and the mutants were transferred into *E. coli* BL21 (DE3) to express the target proteins. Cells with the plasmids were inoculated at 37 °C in Luria Bertani (LB) medium supplemented with $30 \mu\text{g ml}^{-1}$ kanamycin. When OD_{600} value reached 0.6-0.8, 0.5 mM isopropyl β -D-thiogalactopyranoside (IPTG) was

added to induce the expression of target proteins. The induction time for the expression of the wild-type periplasmic domain and the mutant R64A was 6 hours at 26 °C, whereas that for the R73'A, and R69'D mutants was 8 hours at 18 °C. Cells were lysed by sonication in the sonication buffer (50 mM Tris-HCl, pH 8.0 (12), 200 mM NaCl, 2 mM PMSF, 10 mM iminazole). Cell debris was pelleted by centrifugation and the supernatant was applied to 5 ml HisTrap™ HP column (GE Healthcare) equilibrated with buffer A (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 10 mM iminazole). A linear gradient of buffer B (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 500 mM iminazole) was applied to elute the target proteins. The peak fraction containing target proteins from the HisTrap™ HP column was applied to a 120 ml Sephacryl S-200 HR (GE Healthcare) equilibrated with buffer C (50 mM Tris-HCl, pH 8.0, 200 mM NaCl). Proteins were eluted with the same buffer and analyzed by SDS-PAGE.

Design, fabrication, and calibration of the microfluidic device for the chemoeffector selection

Design of the device. A specially designed (unpublished previously) microfluidic device was used in this study. Kim *et al.* recently reported a similar design (13); however, our design is more suitable for large-scale screens for novel chemoeffectors. A schematic representation of the device is shown in Fig. 1A. It consists of a central hole (diameter, 5.0 mm) with 12 circular peripheral holes (diameter, 3.0 mm) around it, connected by means of 12 microchannels to the central hole. Each microchannel is divided into three regions. Regions 1 and 3 have the same dimensions: length 1.5 mm,

width 50 μm , and height 5 μm . Region 2, also termed the analysis region, is 500 μm in length, 200 μm in width, and 25 μm in height. In our design, similar to previously published work (13), agarose gel was used to avoid the convection current and allow the diffusion of small molecules to generate stable linear concentration gradients. But one of the new advantages is that, in our design, the agarose plug can be constructed in either Region 1 or Region 3, so the concentration gradients can be generated either from the peripheral holes to the central hole or from the central hole to the peripheral holes. The two directions of concentration gradients broaden the applications of the device. When agarose plugs were constructed in Region 3, the peripheral holes are the sources for compounds, and the central hole is used as the cells source. Compounds diffused from the peripheral holes to the central hole along the microchannels. This process enabled us to investigate the responses of cells with the same conditions to different chemicals or different chemical concentrations at the same time. When agarose plugs were constructed in Region 1, the central hole is filled with attractant solution, and the peripheral holes are filled with cells of the same or different conditions. Compounds diffused from the central hole to the peripheral holes. This process allowed us to observe the responses of cells under different environmental conditions to the same chemical concentration gradient simultaneously.

Fabrication of the device. Standard soft lithography procedure (14) was used to fabricate microfluidic devices. The well prepared silicon master with the features described above was used to make the mold of microfluidic devices. Polydimethylsiloxane (PDMS, RTV615 044-Pail Kit, crosslinking agent: silicone

potting compound ~1:7, Momentive Specialty Chemicals Inc.) was poured on the master, cured at 75 °C and peeled off. Holes were punched into the patterned PDMS at the positions of the central hole and peripheral holes using cutting tips with tip diameters of 5.0 mm and 3.0 mm (Harris Uni-Core™). The microfluidic devices were bonded to clean microscope cover classes (Fisher Scientific) after treated with oxygen plasma for 1 min in a plasmacleaner (Harrick Plasma) to create hydrophilic devices. Then agarose plugs can be constructed. For the novel chemoeffectors screening, 3% agarose solution incubated at 75 °C was loaded into each peripheral hole at the room temperature. Agarose solution flew into Region 3 and solidified at the entrance of Region 2 (analysis region). For the antagonist function detection, 3% agarose was loaded into the central hole. The agarose solution can flow along Region 1 and stop at the entrance of Region 2. The reasons for the agarose stopped at the interface of Region 2 and Region 1 (Region 3) are described in another study (15). Minimal salt buffer (also termed blank buffer; 10 mM PBS, 0.1 mM EDTA, 0.01 mM L-methionine, 10 mM sodium DL-lactate, pH 7) was loaded into the central hole and the peripheral holes to fill the device with buffer.

Calibration of the device. After the device was well fabricated, fluorescein solution was loaded in the peripheral hole to the final concentration of 100 μM and let it diffuse in wet environment for 15 hours. Then the fluorescence signals in the microchannel were observed using a Nikon Ti-E inverted microscope (Nikon Instruments) with a QuantEM512SC CCD (Roper Scientific). Fluorescence images were recorded using 10× objective lenses. As shown in Fig. S7A, the concentration

gradient of fluorescein is linear in Region 1 and Region 3 at the steady state. The chemical concentration in the microchannel can be described by the one dimensional diffusion equation. At the steady state, the concentration gradient $\nabla C = C_0/l$ is linear, where C_0 is the source concentration of compound, and l is the length of the microchannel (total length of Region 1 and Region 3, 3 mm here). The fluorescent signal in Region 2 is stronger than Region 1 and Region 3, because the height of Region 2 is five times larger. The gradient in Region 2 is not obvious, theoretical explanations are as follows: The equation for the diffusion of compounds is, $\Delta Q' = -D(dc/dx)s\Delta t$, where s is the cross-sectional area, D is the diffusion coefficient. At the steady state, $\Delta Q'$ is the same along microchannel. So in Region 1 (Region 3) and Region 2, $\Delta Q' = -D(dc/dx)_1 s_1 \Delta t = -D(dc/dx)_2 s_2 \Delta t$, $(dc/dx)_1 / (dc/dx)_2 = s_2 / s_1$, where s_1 , s_2 is the cross-sectional area of Region 1 (Region 3) and Region 2 respectively. That is, the gradient is in inverse proportion to s . The cross-sectional area of Region 1 (Region 3) is 19 times larger than that of Region 2, so the gradient in Region 2 is only 1/20 of the gradient in Region 1 (Region 3). Because the concentration change is very small in Region 2, we can estimate that the concentration in Region 2 is almost half of the source concentration of compound, here 50 μM of fluorescein. If we do not consider Region 2, the time t_1 for the concentration gradient to reach a steady state can be approximated $t_1 = l^2/D$, where D is the diffusion coefficient of compounds. It is assumed that the agarose gel has the same diffusion coefficient as water. According to this theoretical estimation, the time for fluorescein to diffuse through the microchannel is about $l^2/D = 5$ h, where $l = 3$

mm (total length of Region 1 and Region 3), $D = 500 \mu\text{m}^2 \text{s}^{-1}$. This is only a simplified method to estimate t_1 , because the different dimension of Region 2 will influence the diffusion time. But 15 hours is sufficient to generate stable linear gradient, as seen in Fig. S7B and C. The linear concentration gradient is stable during the period of experiment, as shown in Fig. S7B and C. The stable time t_2 for the linear concentration gradient can be estimated by the equation $\Delta cV = Ds_1t_2 dc/dl$, where V is the volume of the central hole or the peripheral hole, s_1 is the cross-sectional area of Region 1 or Region 3. If the compounds diffuse along the direction from the peripheral hole to the central hole, at the steady state, the concentration in the central hole is zero. Assuming that after t_2 , the concentration in the central hole changes from zero to $C_0/10,000$, the equation can be rewrote as $t_2 = V_{center}l/10,000Ds_1 = 26.1 \text{ h}$, where $V_{center} = \pi r^2 h \approx 39.25 \text{ mm}^3$ ($h \approx 2 \text{ mm}$), $l = 3 \text{ mm}$, $D = 500 \mu\text{m}^2 \text{s}^{-1}$, $s_1 = 250 \mu\text{m}^2$. If the compounds diffuse along the direction from the central hole to the peripheral hole, the time for the concentration in the peripheral hole changes from zero to $C_0/10,000$ is $t_2 = V_{around}l/10,000Ds_1 = 6.52 \text{ h}$, where $V_{around} \approx 14.13 \text{ mm}^3$. So, the concentration gradients are very stable during the entire period of experiments. This design can have good function even if the concentration changes from zero to $C_0/10$. The time t_2' for the concentration changes from zero to $C_0/10$ is 1,000 times longer than t_2 . So, the device keeps good function during very long period of time. Since a linear gradient was established, the compound concentration range in Region 1 was about 0–50% of the source concentration in the peripheral hole.

Cell and compound preparation for the microfluidic experiments to

select novel chemoeffectors. Single colonies of *E. coli* strains RP437, UU1624, and RP2361 expressing GFP proteins were grown at 30 °C overnight in Tryptone Broth medium (TB, 10 g L⁻¹ tryptone and 5 g L⁻¹ NaCl) supplemented with 100 µg ml⁻¹ ampicillin. The grown cultures were then diluted with 100 times by fresh TB medium containing antibiotics and grown at 30 °C until OD₆₀₀ had reached ~0.3. Cells were harvested by centrifugation at 3,000 rpm for 5 min at room temperature. The supernatant was discarded and the pelleted cells were washed twice with minimal salt buffer to remove remaining TB medium (16-17). Finally, cells were resuspended in minimal salt buffer. All the compounds used in the microfluidic experiments were dissolved in minimal salt buffer (pH7).

Microfluidic experiments to detect cell responses to L-arginine.

The cell preparation was similar as described above, except that the *E. coli* strains UU1250 expressing wild-type or mutant Tar receptor were grown in TB medium supplemented with 100 µg ml⁻¹ ampicillin, 30 µg ml⁻¹ chloramphenicol and 500 µM IPTG. The responses of *E. coli* strains expressing wild-type Tar, R69'E, R73'E or R69'ER73'E mutant Tar to the gradients of arginine were measured by microfluidics. Since we only need to detect response to a single ligand (Arg), we did not use the multi-channel radial design developed for parallel compound screening shown in Fig. 1A. Instead, a previously reported simpler microfluidic device was used (15). This microfluidic device is more suitable for detecting the responses of different strains to the same compound simultaneously. The design, fabrication, and calibration of this microfluidic are described in detail in (15). Experiments were operated as described in

(15). In short, a sink side pore and a source side pore are connected by an observation channel and agarose gel channels. The prepared *E. coli* cells were added in the sink side pores of the microfluidic device. It takes roughly one hour for the cells to diffuse into the observation channel and reach a steady state. Then, we added compound solutions in the source side pores. The compound will diffuse into the observation channel and establish a concentration gradient gradually. After adding the compound, images were recorded to detect the cell fluorescent intensities in the observation channel. The images were captured every 5 min for 90 min. The responses of cells were characterized by the fluorescence intensities in the analysis region (yellow rectangle) of the observation channel (Fig. 6A). Data were analyzed by Image J.

Expression level analysis using Western Blot. The expression level of mutant Tar receptors expressed from pPD12 derivatives were measured in UU1250 using Western Blot. Strains were grown and suspended in minimal salt buffer as described above. Cells were lysed by boiling and subjected to electrophoretic separation using SDS-PAGE. Proteins were transferred from the gel to the nitrocellulose membrane, treated by rabbit polyclonal anti-Tar antibody and detected by goat anti-rabbit IgG (AP) secondary antibody. Intensity profiles in individual lanes were analyzed using Image J. The relative amounts of Tar proteins in different lanes were compared by using a chromosomally encoded protein as the internal standard. The expression level of Tar mutant R69'E, R73'E or R69'ER73'E was compared with that of wild-type Tar (expressed from pPD12)

Statistical analysis. Statistical analysis was done using GraphPad Prism 5.04 (GraphPad Software). For statistical comparison, one-way ANOVA followed by *Dunnnett's* test, or a Student's two-tailed unpaired *t*-test were used. $P < 0.05$ was considered significant.

SI References

1. Tellinghuisen J (2008) Isothermal titration calorimetry at very low *c*. *Anal Biochem* 373:395–397.
2. Turnbull WB, Daranas AH (2003) On the value of *c*: can low affinity systems be studied by Isothermal titration calorimetry? *J Am Chem soc* 125: 14859-14866
3. Sager BM, Sekelsky JJ, Matsumura P, Adler J (1988) Use of a computer to assay motility in bacteria. *Anal Biochem* 173(2):271-277.
4. Milburn MV, *et al.* (1991) Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* 254(5036):1342-1347.
5. Yeh JI, Biemann HP, Pandit J, Koshland DE, Kim SH (1993) The three-dimensional structure of the ligand-binding domain of a wild-type bacterial chemotaxis receptor. Structural comparison to the cross-linked mutant forms and conformational changes upon ligand binding. *J Biol Chem* 268(13):9787-9792.
6. Yeh JI, *et al.* (1996) High-resolution structures of the ligand binding domain

- of the wild-type bacterial aspartate receptor. *J Mol Biol* 262(2):186-201.
7. Chi YI, Yokota H, Kim SH (1997) Apo structure of the ligand-binding domain of aspartate receptor from *Escherichia coli* and its comparison with ligand-bound or pseudoligand-bound structures. *FEBS Lett* 414(2):327-332.
 8. Bowie JU, Pakula AA, Simon MI (1995) The three-dimensional structure of the aspartate receptor from *Escherichia coli*. *Acta Crystallogr D Biol Crystallogr* 51(Pt 2):145-154.
 9. Xiang Z, Honig B (2001) Extending the accuracy limits of prediction for side-chain conformations. *J Mol Biol* 311(2):421-430.
 10. Brooks BR, *et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30(10):1545-1614.
 11. Huey R, Morris GM, Olson AJ, Goodsell DS (2007) A semiempirical free energy force field with charge-based desolvation. *J Comput Chem* 28(6):1145-1152.
 12. Milligan DL, Koshland DE, Jr. (1993) Purification and characterization of the periplasmic domain of the aspartate chemoreceptor. *J Biol Chem* 268(27):19991-19997.
 13. Kim M, Kim T (2010) Diffusion-based and long-range concentration gradients of multiple chemicals for bacterial chemotaxis assays. *Anal Chem* 82(22):9401-9409.
 14. Yang T, Jung S, Mao H, Cremer PS (2001) Fabrication of phospholipid bilayer-coated microchannels for on-chip immunoassays. *Anal Chem*

73(2):165-169.

15. Si G, Yang W, Bi S, Luo C, Ouyang Q (2012) A parallel diffusion-based microfluidic device for bacterial chemotaxis analysis. *Lab Chip* 12(7):1389-1394.
16. Kalinin YV, Jiang L, Tu Y, Wu M (2009) Logarithmic sensing in *Escherichia coli* bacterial chemotaxis. *Biophys J* 96(6):2439-2448.
17. Mao H, Cremer PS, Manson MD (2003) A sensitive, versatile microfluidic assay for bacterial chemotaxis. *Proc Natl Acad Sci USA* 100(9):5449-5454.
18. Ames P, Studdert CA, Reiser RH, Parkinson JS (2002) Collaborative signaling by mixed chemoreceptor teams in *Escherichia coli*. *Proc Natl Acad Sci USA* 99(10):7060-7065.
19. Cormack BP, Valdivia RH, Falkow S (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173(1 Spec No):33-38.
20. Hansen MC, Palmer RJ, Jr., Udsen C, White DC, Molin S (2001) Assessment of GFP fluorescence in cells of *Streptococcus gordonii* under conditions of low pH and low oxygen concentration. *Microbiology* 147(Pt 5):1383-1391.
21. Sourjik V, Berg HC (2004) Functional interactions between receptors in bacterial chemotaxis. *Nature* 428(6981):437-441.
22. Yang Y, Sourjik V (2012) Opposite responses by different chemoreceptors set a tunable preference point in *Escherichia coli* pH taxis. *Mol Microbiol* 86(6):1482-1489.
23. Derr P, Boder E, Goulian M (2006) Changing the specificity of a bacterial

- chemoreceptor. *J Mol Biol* 355(5):923-932.
24. Parkinson JS, Houts SE (1982) Isolation and behavior of *Escherichia coli* deletion mutants lacking chemotaxis functions. *J Bacteriol* 151(1):106-113.
 25. Bibikov SI, Miller AC, Gosink KK, Parkinson JS (2004) Methylation-independent aerotaxis mediated by the *Escherichia coli* Aer protein. *J Bacteriol* 186(12):3730-3737.
 26. Gosink KK, Buron-Barral MC, Parkinson JS (2006) Signaling interactions between the aerotaxis transducer Aer and heterologous chemoreceptors in *Escherichia coli*. *J Bacteriol* 188(10):3487-3493.
 27. Manson MD, Blank V, Brade G, Higgins CF (1986) Peptide chemotaxis in *E. coli* involves the Tap signal transducer and the dipeptide permease. *Nature* 321(6067):253-256
 28. Slocum MK, Parkinson JS (1985) Genetics of methyl-accepting chemotaxis proteins in *Escherichia coli*: null phenotypes of the *tar* and *tap* genes. *J Bacteriol* 163(2):586-594.

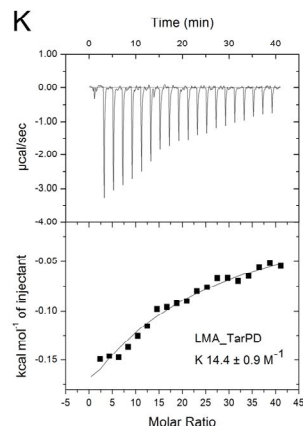
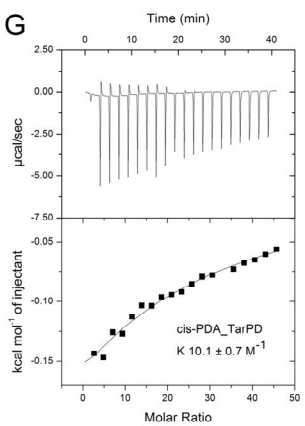
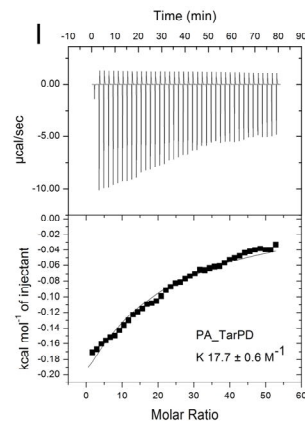
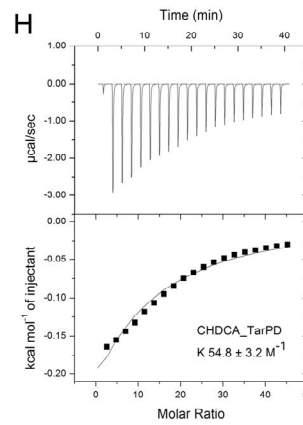
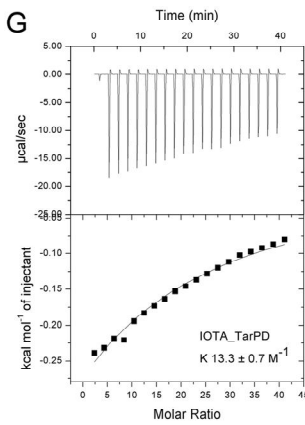
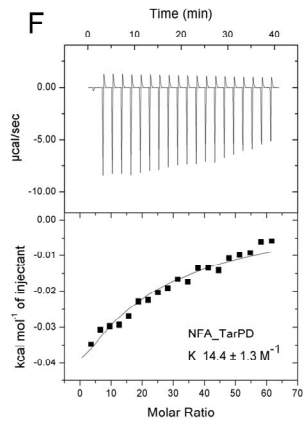
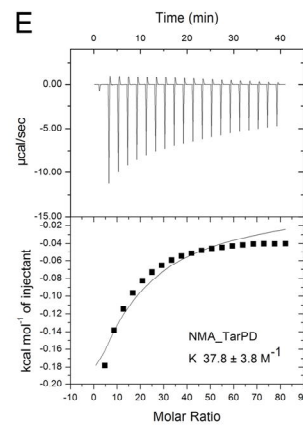
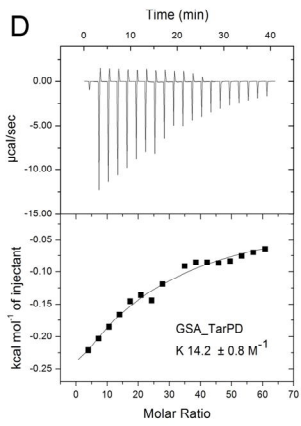
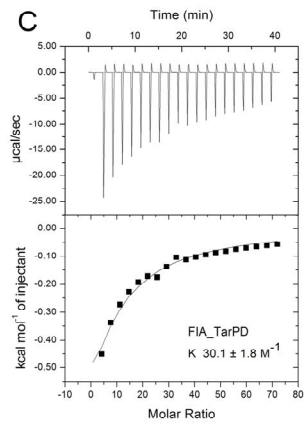
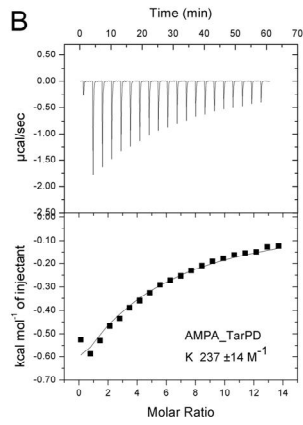
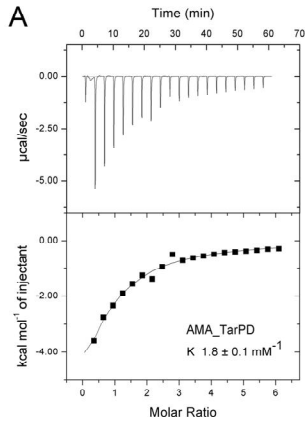


Fig. S1. Binding affinity measurements by ITC. Binding affinity of Tar periplasmic domain with (A) AMA, (B) AMPA, (C) FIA, (D) GSA, (E) NMA, (F) NFA, (G) IOTA, (H) CHDCA, (I) PA, (J) *cis*-PDA, and (K) LMA.

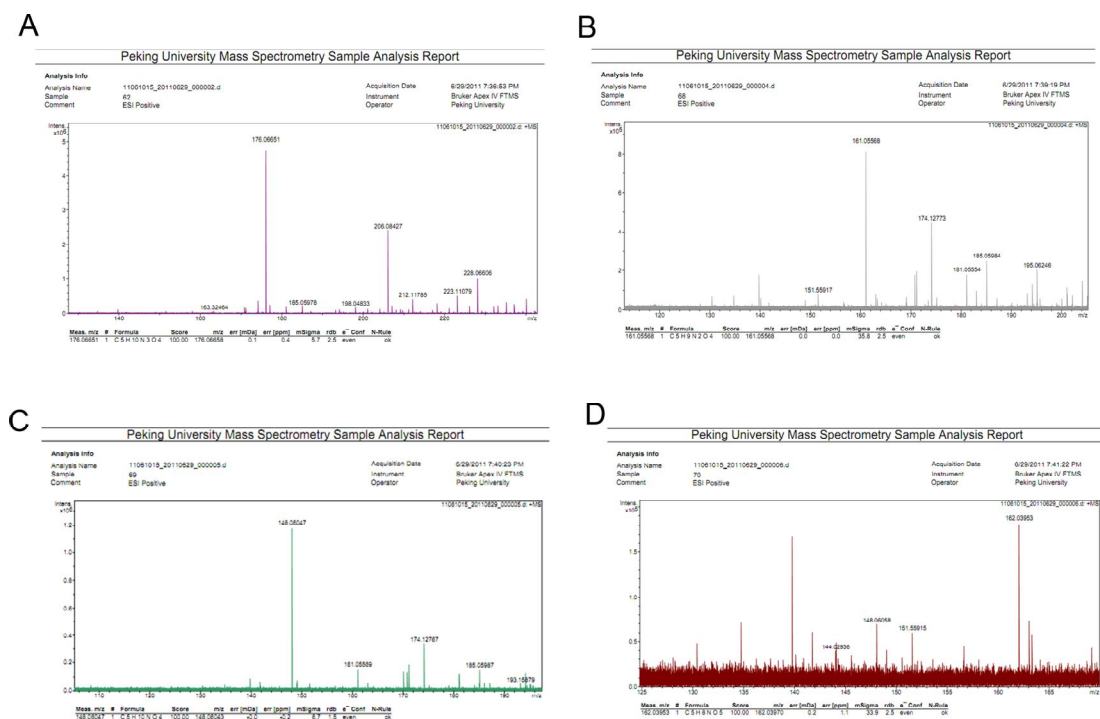


Fig. S2. Results of ESI⁺ experiments for the novel attractants (A) GSA, (B) FIA, (C) NMA, and (D) NFA. The molecular weight for L-aspartate should be: $M = 133.10$, $M + H = 134.10$, $M + Na^+ = 156.10$. There is no peak at those three positions, so the four attractants do not contain aspartate contamination.

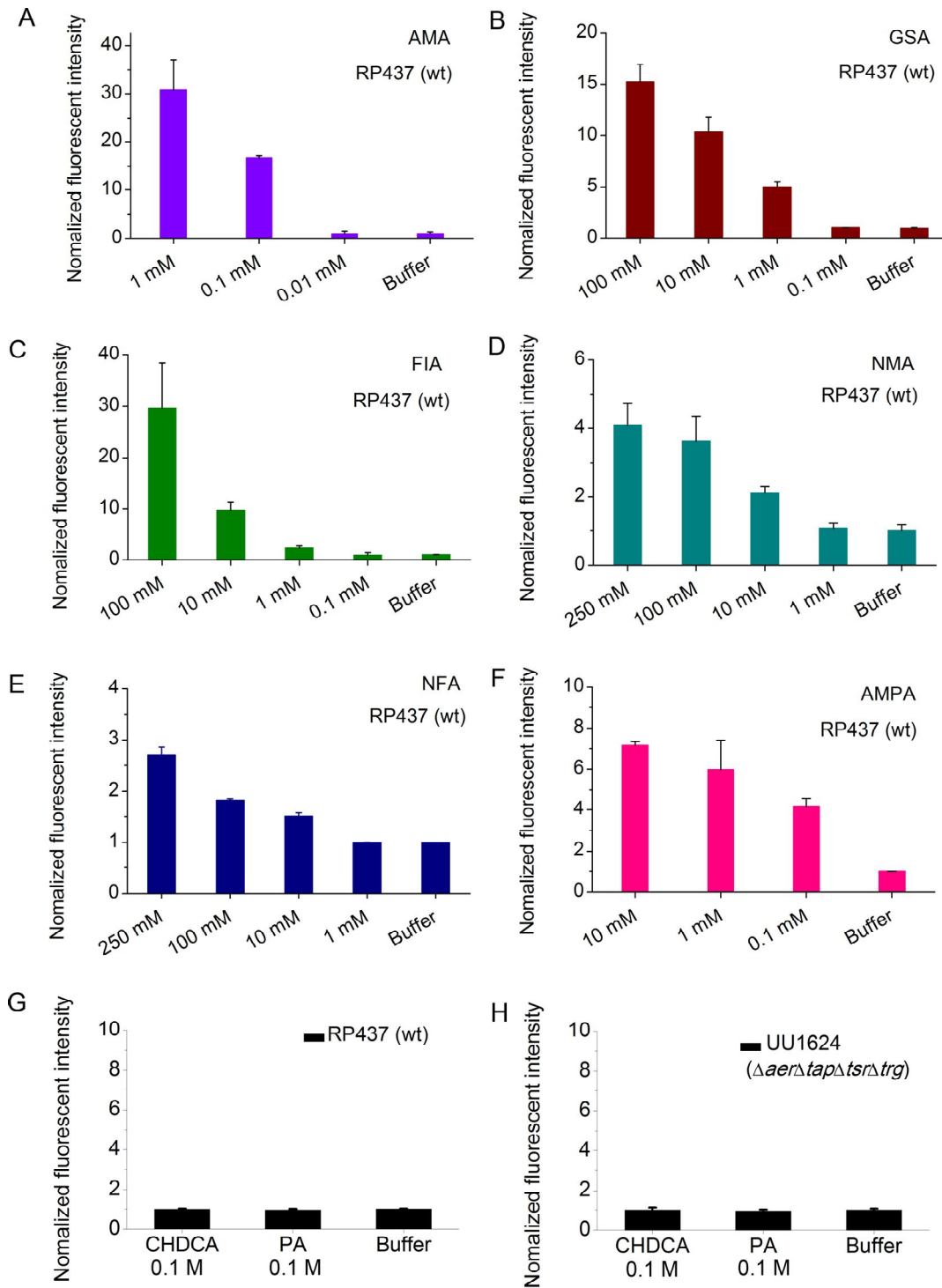


Fig. S3. Responses of *E. coli* cells to different source concentrations of novel attractants, CHDCA and PA. RP437 responses to different source concentrations of (A) AMA, (B) GSA, (C) FIA, (D) NMA, (E) NFA and (F) AMPA were measured in

the analysis region (Region 2) (mean \pm SD, $n = 2$). CHDCA and PA cannot attract (*G*) RP437 and (*H*) UU1624 even if the source concentrations were in the order of 0.1 M. The fluorescent intensities in the analysis region for these two compounds were similar with that in the blank buffer (mean \pm SD, $n = 2$). The cell concentration and exposure time were determined by each experiment.

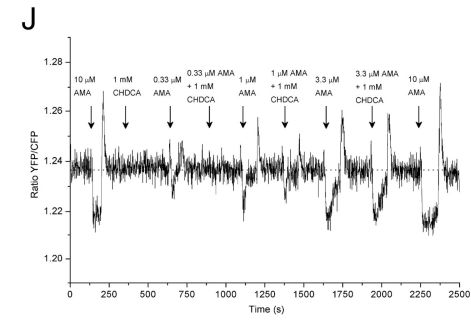
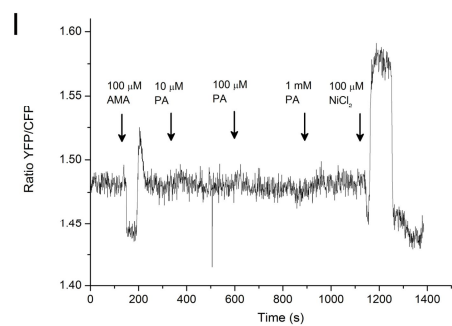
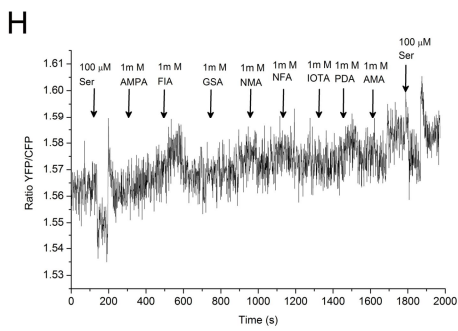
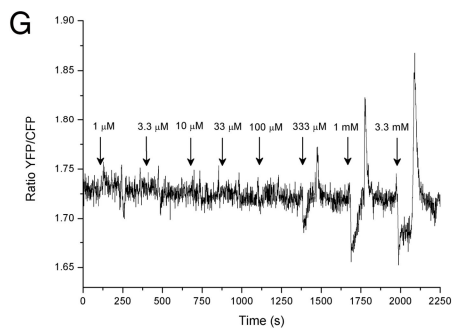
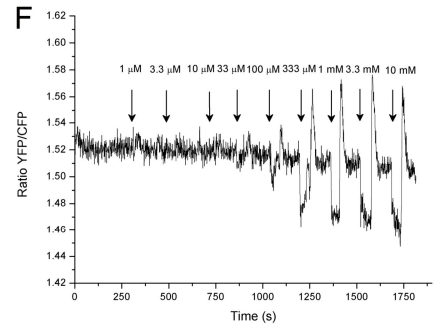
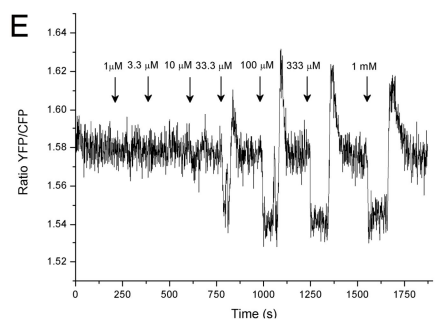
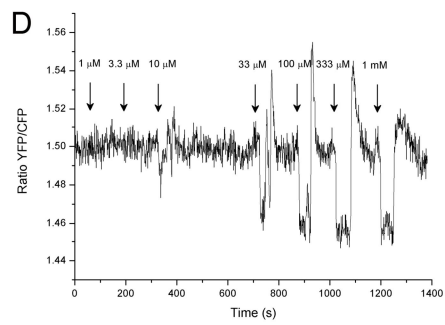
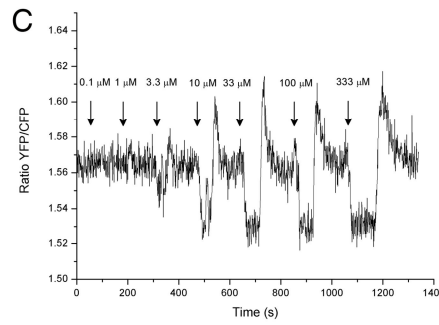
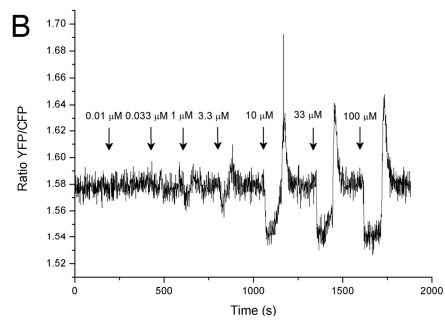
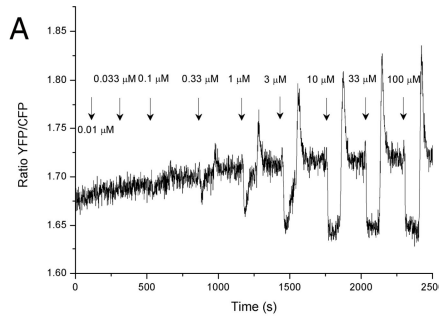


Fig. S4. Intracellular responses of *E. coli* to novel chemoeffectors observed by FRET. The Tar-only strain that expresses CheY-YFP and CheZ-CFP pair were stimulated by stepwise addition or removal of attractants (A) AMA, (B) AMPA, (C) FIA, (D) GSA, (E) NMA, (F) NFA, (G) IOTA. (H) The Tsar-only strain cannot response to novel attractants. (I) PA cannot induce the change of FRET signals. (J) FRET measurement of the influence of CHDCA on response to steps of AMA.

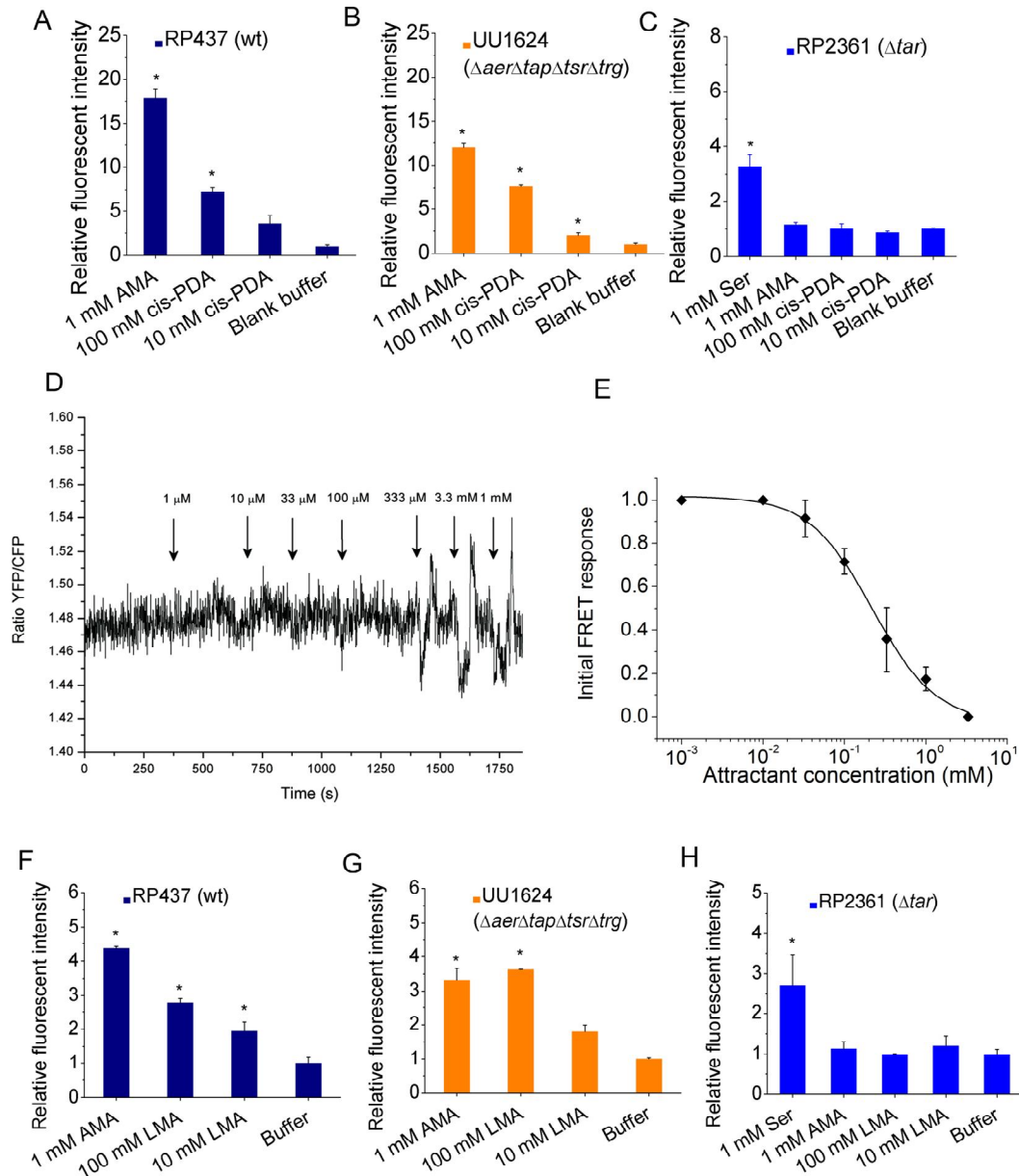


Fig. S5. Responses of *E. coli* to *cis*-PDA and LMA. Relative fluorescence intensities in the analysis region emitted by stratins (A) RP437, (B) UU1624, and (C) RP2361 responding to different source concentrations of *cis*-PDA (mean \pm SEM, $n = 3$). Cell responses to buffer were set to one. Strains expressing a functional Tar receptor are attracted by *cis*-PDA in the microfluidic experiments. (D) FRET measurement, plotted as a change in YFP/CFP ratio, to stepwise addition or

subsequent removal of *cis*-PDA. (E) Dose-response curve for *cis*-PDA calculated from (D). Relative fluorescence intensities in the analysis region emitted by strains (F) RP437, (G) UU1624, and (H) RP2361 in response to different source concentrations of LMA were recorded (mean \pm SEM, $n = 2$). Strains expressing functional Tar receptors are attracted by LMA. Asterisks in (A)-(B), and (F)-(H) indicate statistical significance ($P < 0.05$) compared to the blank buffer by one-way ANOVA.

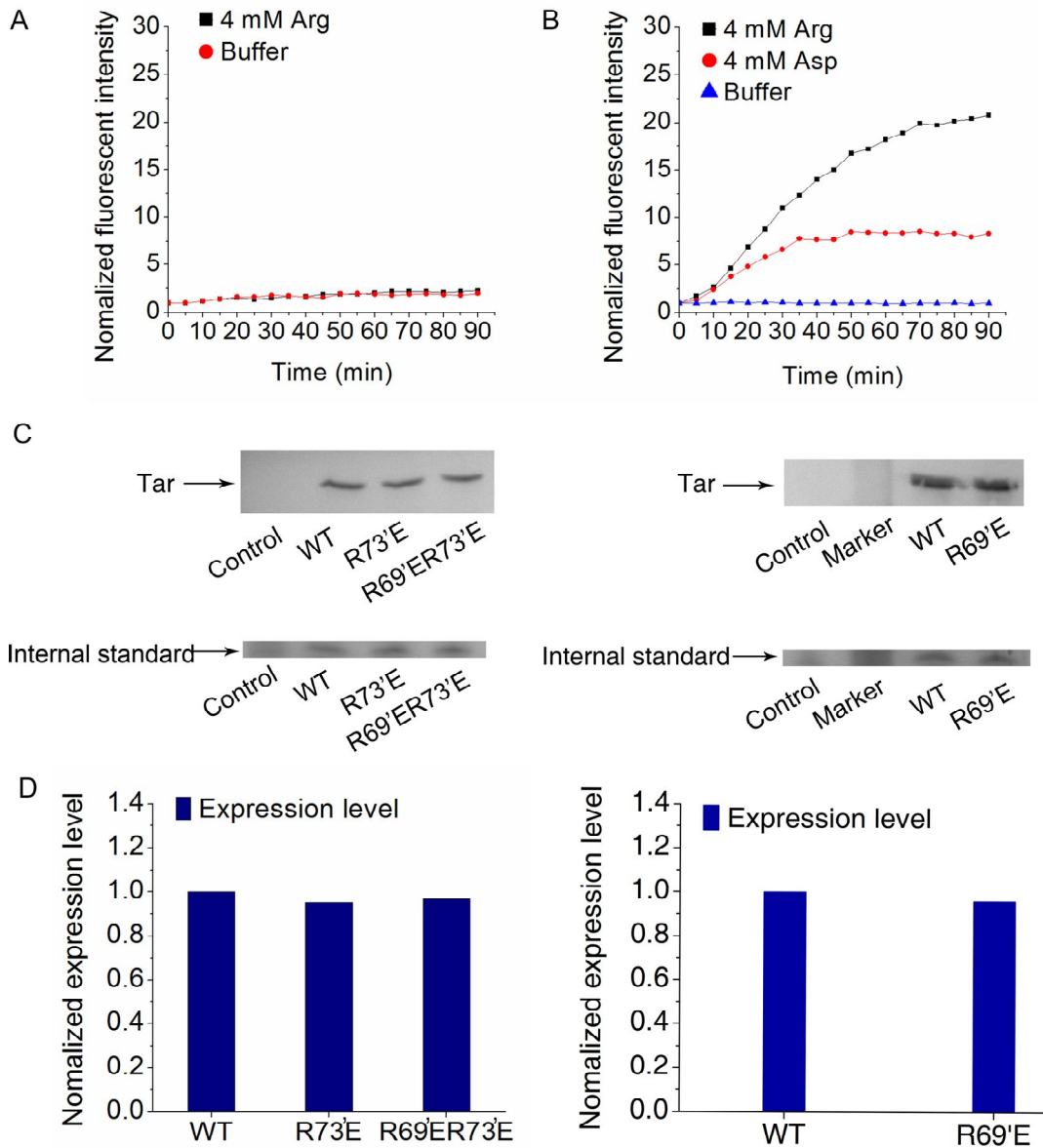


Fig. S6. (A) The *E. coli* strain with wild-type Tar cannot sense L-arginine as an attractant. (B) The Tar mutant R69'ER73'E has stronger chemotaxis response to L-arginine than L-aspartate. (C) The expression level of wild-type and mutant Tar R69'E, R73'E and R69'ER73'E determined by Western Blot. A chromosomally encoded protein was used as the internal standard. UU1250 with pPD10 plasmid was the control. (D) R69'E, R73'E and R69'ER73'E have similar expression level with the

wild-type Tar (WT). The concentrations given in (A) and (B) are the maximum concentrations at the right end of the observation channel.

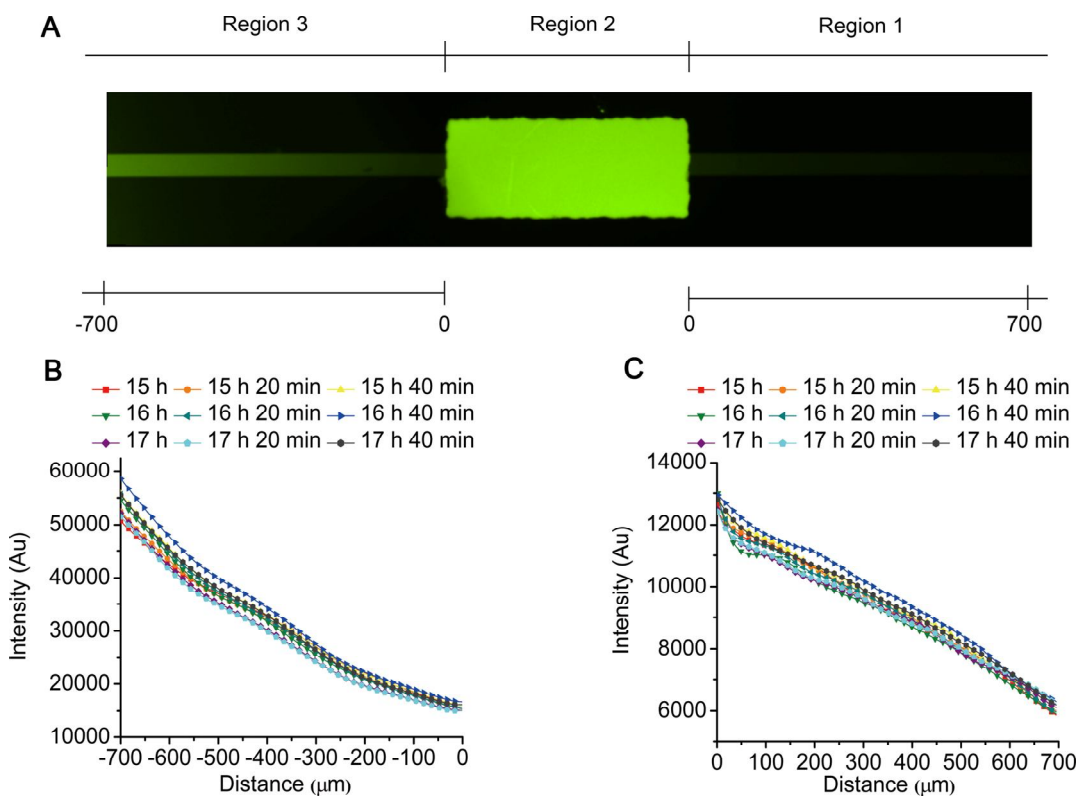


Fig. S7. Microfluidic device calibration. (A) The profile of the concentration gradient of fluorescein imaged in the microchannel after diffusing for 15 hours. The source concentration for the fluorescein is $100 \mu\text{M}$. After 15 hours, the concentration gradient maintains linear and stable during the period of experiments. The coordinate of X-axis (B) -700-0 and (C) 0-700 correspond to the coordinate signed in (A).

Table S1 Results of ITC

	$[M_0]^*$ (mM)	$[X_0]^\dagger$ (mM)	K_a (M^{-1})	K_d (mM)	c^\ddagger	Saturation (%)
AMA	0.34	10	1790 ± 116	0.559 ± 0.036	0.304	95
AMPA	0.3	20	237 ± 14	4.2 ± 0.2	0.036	83
FIA	1	350	30.1 ± 1.8	33 ± 2	0.015	91
GSA	0.85	300	14.2 ± 0.8	70 ± 4	0.006	81
NMA	0.5	200	37.8 ± 3.8	26 ± 3	0.009	88
NFA	1	300	14.4 ± 1.3	69 ± 6	0.007	81
IOTA	1	200	13.3 ± 0.7	75 ± 4	0.007	73
CHDCA	0.5	110	54.8 ± 3.2	18 ± 1	0.014	86
PA	1	250	17.7 ± 0.6	56 ± 2	0.009	82
<i>cis</i> -PDA	1	200	10.1 ± 0.7	99 ± 7	0.005	67
LMA	1	200	14.4 ± 0.9	69 ± 4	0.007	74

* $[M_0]$ is the concentration of proteins

† $[X_0]$ is the concentration of compounds

‡ $c = nK_a[M_0]$, $n = 0.5$

Table S2 Mean speed and mean angular speed analysis of UU1624 cells swimming in the ambient blank buffer and 1mM CHDCA, pH 7

Swimming time	5 min		60 min	
	Blank buffer	1mM CHDCA	Blank buffer	1mM CHDCA
Swimming medium				
Number of cells tracked	26	31	27	24
Tracking time (s)	20	20	20	20
Mean speed ($\mu\text{m s}^{-1}$) ^{*,†}	17 ± 6	17 ± 7	14 ± 3	15 ± 6
Mean angular speed (deg frame ⁻¹) ^{*,†}	53 ± 16	51 ± 15	47 ± 17	44 ± 13

* mean ± SD

† Frame interval: 0.07 s

Table S3 Strains and plasmids used in this study

	Genotype or phenotype	Description	Source or reference
Plasmids			
pLC113	<i>tar</i> Cam ^r	Expresses <i>E. coli</i> wild-type full-length Tar receptor	(18)
pMDL101	<i>tar (33-188)</i> Kan ^r	Expresses <i>E. coli</i> wild-type Tar periplasmic domain, N-terminal His-tag	This study
GFP plasmid	<i>gfpmut2</i> Amp ^r	Expresses GFP proteins	(19)
pCM18	<i>gfpmut3</i> Cam ^r	Expresses GFP proteins	(20)
pVS88	<i>cheZ-ecfp / cheY-eyfp</i> Amp ^r	Expresses FRET pair CheY-YFP and CheZ-CFP	(21)
pVS1092	<i>tar [QEQE]</i> Cam ^r	Expresses <i>E. coli</i> Tar receptor	(22)
pVS1252	<i>tsar</i> Cam ^r	Expresses T _{sar} receptor, the chimera receptor replacing the periplasmic domain of Tar with the periplasmic domain of T _{sar}	(22)
pPD12	<i>tar</i> Amp ^r	Expresses <i>E. coli</i> wild-type full-length Tar receptor. Used for mutagenesis for Tar rational design. <i>tar</i> was ligated into pPD10.	(23)
pPD10	Amp ^r	Expression plasmid	(23)
<i>E. coli</i> strains			
BL21 (DE3)	F ⁻ , <i>ompT hsdS_B (r_B⁻ m_B⁻) gal dcm</i> (DE3)	Periplasmic domain of Tar expression and purification strain	Novagen, Germany
RP437	<i>thr-1leuB6 his-4 metF59 eda-50 rpsL136</i>	Wild-type <i>E. coli</i> strain	(24)

UU1250	$\Delta aer-1 \Delta(tar-tap)5201 \Delta tsr-7028$ $\Delta trg-100$ <i>ygjG::Gm zbd::Tn5</i>	The strain lacks all five chemoreceptors Aer, Tar, Tsr, Trg, and Tap	(25)
UU1624	$\Delta aer-1 \Delta tap-3654$ $\Delta tsr-7028 \Delta trg-100$ <i>ygjG::Gm</i> <i>zbd::Tn5</i>	The strain possesses only the Tar chemoreceptor	(26)
RP2361	$\Delta tar-3862$	The strain lacks the chemoreceptor Tar	(27-28)
VS181	$\Delta(ch eY cheZ) \Delta tsr \Delta(tar tap)$ $\Delta trg \Delta aer$	The strain lacks all five chemoreceptors Aer, Tar, Tsr, Trg, Tap, as well as CheY and CheZ.	(21)
Tar-only strain	VS181 <i>tar</i>	VS181 with pVS88 and pVS1092. Used in FRET measurement.	(22)
Tsar-only strain	VS181 <i>tsar</i>	VS181 with pVS88 and pVS1252. Used in FRET measurement.	(22)
