






Tailoring *Escherichia coli* Chemotactic Sensing towards Cadmium by Computational Redesign of Ribose-Binding Protein

 Hengyi Li,^a  Changsheng Zhang,^a Xi Chen,^b Hantian You,^a  Luhua Lai^{a,b,c}

^aBNLMS, College of Chemistry and Molecular Engineering, Peking University, Beijing, China

^bCenter for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China

^cPeking-Tsinghua Center for Life Sciences, Peking University, Beijing, China

Hengyi Li and Changsheng Zhang contributed equally to this article.

ABSTRACT Periplasmic binding proteins such as ribose-binding proteins (RBPs) are involved in the bacterial chemotaxis two-component system. RBP selectively identifies and interacts with ribose to induce a conformational change that leads to chemotaxis. Here, we report the development of an engineered *Escherichia coli* (*E. coli*) strain expressing a redesigned RBP that can effectively sense cadmium ions and regulate chemotactic movement of cells toward a cadmium ion gradient. RBP was computationally redesigned to bind cadmium ions and produce the conformational change required for chemoreceptor binding. The successful design, CdRBP1, binds to cadmium ions with a dissociation constant of 268 nM. When CdRBP1 was expressed in the periplasmic space of *E. coli*, the bacteria became live cadmium ion hunters with high selectivity over other divalent metal ions. This work presents an example of making cadmium ions, which are toxic for most organisms, as an attractant to regulate cells movement. Our approach also demonstrates that RBP can be precisely designed to develop metal-detecting living systems for potential applications in synthetic biology and environmental studies.

IMPORTANCE Cadmium pollution is one of the major environmental problems due to excessive release and accumulation. New technologies that can auto-detect cadmium ions with good biocompatibility are in urgent need. In this study, we engineered the bacterial chemotaxis system to positively sense cadmium ions by redesigning ribose-binding protein (RBP) to tightly bind cadmium ion and produce the right conformational change for receptor binding and signaling. Our engineered *E. coli* cells can auto-detect and chase cadmium ions with divalent metal ion selectivity. Many attempts have been carried out to redesign RBP at the ribose binding site with little success. Instead of the ribose binding site, we introduced the cadmium binding site in the opening of the ribose binding pocket by a specially developed computational algorithm. Our design strategy can be applied to engineer live bacteria with autonomous detection and remediation abilities for metal ions or other chemicals in the future.

KEYWORDS bacterial chemotaxis, cadmium detection, computational design, protein engineering, ribose-binding protein

Bacteria rely on two-component systems to perceive and transduce environmental stimuli (1), including the chemotaxis system. This two-component chemotaxis system can accurately control cell motility in response to changes in environmental chemo-effectors (1–3). Cells with the chemotaxis system can efficiently and rapidly approach chemically favorable environments, such as those with higher concentrations of sugars, amino acids, and peptides, and avoid unfavorable ones, such as those

Editor Danielle Tullman-Ercek, Northwestern University

Copyright © 2022 Li et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Luhua Lai, lhilai@pku.edu.cn.

The authors declare no conflict of interest.

Received 30 August 2021

Accepted 13 December 2021

Published 11 January 2022

consisting of ions and acids (4, 5). The ability of bacteria to sense and respond to novel effectors in their surroundings and to direct their migration would be useful in bioremediation (6, 7), biomedical deliveries (8) and biosensor developments (9, 10).

Periplasmic binding proteins (PBPs) are a versatile superfamily of proteins involved in mediating bacterial chemotaxis (11). PBPs have been shown to interact with a variety of carbohydrates, anions, dipeptides, and amino acids that act as either attractant or repellent molecules (12). However, metal ions functioning as attractants through PBP binding have yet to be reported. Although PBPs bind a broad array of ligands, their three-dimensional structural folds are highly conserved. PBPs consist of two domains connected by a hinge region, with the ligand-binding site located at the interface of the two domains (13). Ligand binding elicits a conformational change that results in the formation of a protein-binding surface for chemoreceptor recognition (14, 15). Recently, a ligand was reported binding to the PBP, periplasmic glucose (or galactose) binding protein, and preventing the domain closure of this PBP. As a result, it acted as an antagonist that inhibits *E. coli* glucose chemotaxis (16).

Both the diversity of PBPs and their relatively flexible protein structures are helpful for the diverse chemotaxis ligands. Therefore, engineering of PBPs may give bacteria the ability to sense and respond to new effectors. Several attempts to construct PBP mutants that bind novel ligands at the ligand-binding sites have been reported (17, 18). However, most of these designs resulted in unfavorable protein folds or undesired binding abilities. Designing PBPs capable of recognizing and directing cell migration toward novel ligands still remains challenge.

Cadmium ion (Cd^{2+}), as one of the heavy metals, is a significant threat to the environment due to accidental release. Strategies for cadmium detection and remediation in the environment have been developed (19–24). Emerging methods based on microorganisms for *in situ* bioremediation and biodetection of cadmium offer an alternative means for rapid, on-site heavy metal detection and remediation (25, 26). However, the movement of microorganisms is often hindered by the toxic nature of Cd^{2+} , which results in inadequate distribution of cells in pollutant regions (25). Because the microorganism cells cannot actively seek cadmium, leading cadmium targeting ratio is still very low. Therefore, constructing bacterial chemotaxis behavior that could serve as self-guided motors to actively seek cadmium would be a preferable way to address these issues.

In the present study, we generated an *E. coli* strain that can recognize and make a positive chemotaxis response to Cd^{2+} , by engineering ribose-binding protein (RBP), which is one of the PBPs in *E. coli*. RBP selectively identifies and interacts with ribose (15, 27), which induces a change in protein conformation from an open to a closed state. The closed state of RBP allows further interaction with and activation of the Trg chemoreceptor (2, 28). Residues in RBP that are essential for chemotactic function (29), structural stability (30), periplasmic abundance, or signaling (31) have been identified. Herein, we developed a computational design algorithm systematically and efficiently searching for Cd^{2+} -binding sites at the domain–domain interface in the RBP closed structure, in order to discover RBP mutants with the behavior of Cd^{2+} -binding induced domain closure. With a series of experimental studies, we identified CdRBP1 as an RBP mutant that binds Cd^{2+} tightly and maintains the closed state conformation that is required for chemotaxis signaling. The engineered *E. coli* strain that expresses CdRBP1 in the periplasmic space exhibited positive chemotactic movement toward Cd^{2+} (Fig. 1).

RESULTS

The computational design strategy. In a previous study, we have developed a computational algorithm, URANTEIN, to search for scaffold proteins for pockets that may accommodate uranyl ions and to subsequently design the uranyl binding site in the selected scaffold protein (32). For the selected scaffold protein, a library collecting the positions and conformations of all possible coordination residues was built first, and then a depth-first set-reduction algorithm (32, 33) was used for searching the

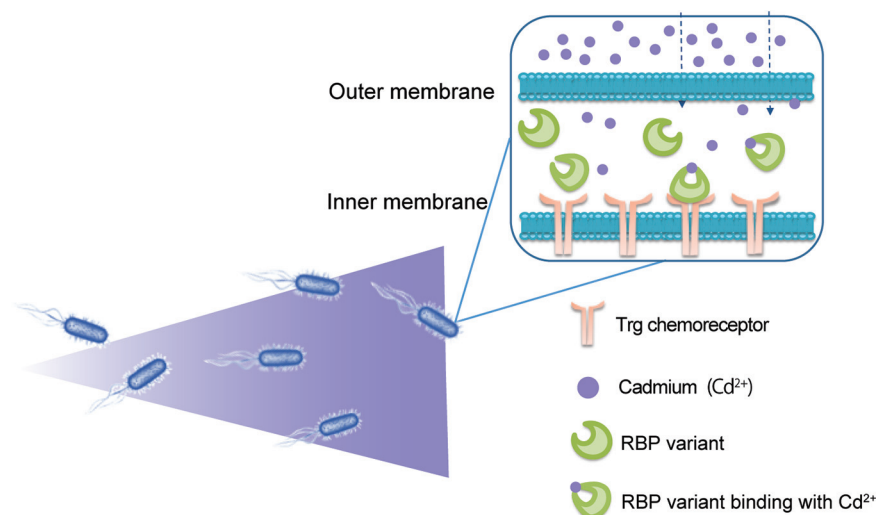


FIG 1 Schematic illustration of RBP variant-based *E. coli* chemotaxis mechanism. After the RBP variants expressed in the *E. coli* periplasmic region bind to cadmium ions, they undergo a conformational change from the open to the closed state that can interact with the Trg transmembrane receptor. This conformational change activates the signaling cascade that promotes *E. coli* chemotaxis toward cadmium ions.

combinations of these residues for the strong binding with uranyl ions. Using a similar strategy, in the present study we developed a computational algorithm for designing Cd^{2+} binding sites on a given scaffold protein. The corresponding program, CADEIN, can be accessed at: <https://github.com/victorPKU/CADEIN>. The computational process is illustrated in Fig. 2. The following is a general description of the method, and more details are given in the Materials and Methods section. Sulfur atoms in the cysteine side chain, carboxylate oxygen atoms in glutamate and aspartate, nitrogen atoms in the histidine side chain, and backbone carbonyl oxygen atoms are used as potential Cd^{2+} coordinating atoms according to the statistical analysis of the Cd^{2+} -binding proteins in the Protein Data Bank (34).

In the present study, the closed structure of RBP (Protein Data Bank code 2DRI) was used as the protein scaffold (35), and the Cd^{2+} binding site was designed at the interface between the two domains. Each interfacial residue was computationally mutated to one of the Cd^{2+} coordination residues, cysteine, glutamate, aspartate, and histidine, using the side chain assembly method and the corresponding backbone-dependent rotamer libraries (36). All of these possible coordination residues at candidate sites formed the coordination residue library. Then, the positions for Cd^{2+} binding of each residue in the library were set. Coordination residues with Cd^{2+} -binding positions close to each other (within 1.6 Å) were combined. All of the coordination pairs were identified and formed a coordination pair library. Based on the connectivity information of coordination residues in the pair library, a set-reduction algorithm (32, 33), which used a depth-first search procedure for the successive elimination of candidate combinations of coordination residues, was used to get the multi-residue-coordination solutions. To obtain high binding affinity mutants, we limited the search to solutions with four or five coordination residues containing at least two cysteines. Five high resolution crystal structures of Cd^{2+} -binding proteins (resolution ≤ 1.5 Å, PDB codes: 5AI3 (37), 1R0I (38), 4CVS (39), 2OA9 (40), and 1WB4 (41)) were used as references for analyzing the Cd^{2+} coordination geometry and designing the scoring function. The average values of coordination bonds, bond angles, and dihedral angles (shown in Figure S1a in the supplemental material) in these structures were used as standard parameters and are presented in Table S1. If all these cadmium coordination geometry parameters of a solution meet the “acceptable” thresholds (Table S1), then the solution was retained. If all the parameters of a coordination residue in this solution meet the

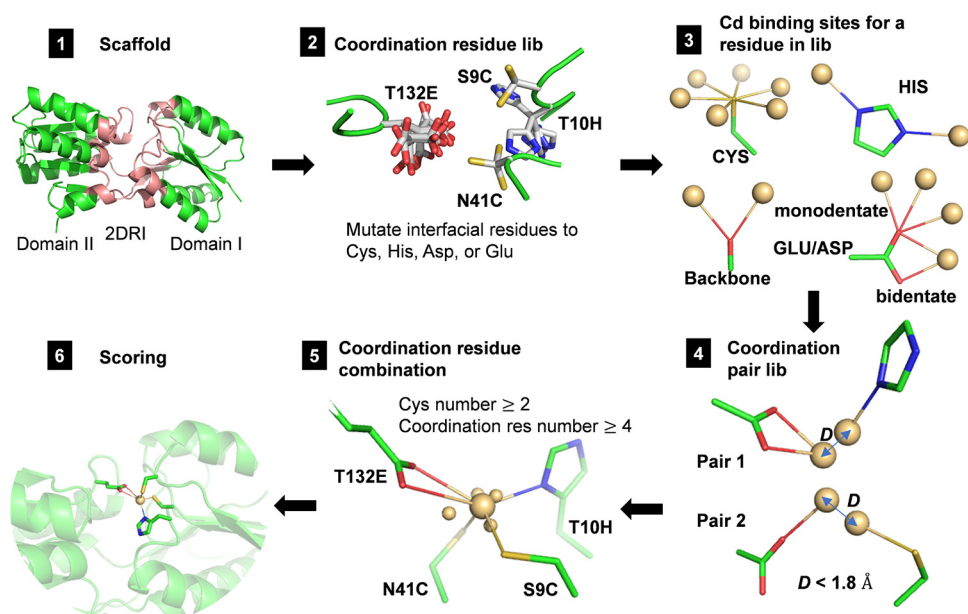


FIG 2 Computational process for designing Cd²⁺-binding ribose-binding protein (RBP) mutants. (1) The structure of the RBP in the holo state, 2DR1, was used as the starting scaffold protein. Only residues at the domain-domain interface (colored in pink) were mutable for Cd²⁺ coordination. (2) All of the mutable residues were mutated to Cd²⁺ coordination residues, including cysteine, histidine, aspartate, and glutamate, using a side chain assembly method. If the assembled side chain did not exhibit steric hindrance with surrounding residues, it was saved in the coordination residue library. The backbone oxygens were also saved to this library. (3) In this figure, T132E, S9C, T10H, and N41C are shown as examples from the coordination residue library. For each coordination atom in the residue library, Cd²⁺ binding position(s) were generated using a set of standard geometry parameters (Table S1, column 3). For the carboxyl group of aspartate or glutamate, one bidentate and three monodentate positions were generated. (4) If the distance (denoted as *D* in the figure) between two residues (at the standard coordination site) in the coordination library was less than 1.8 Å, and the two residues did not interfere with each other sterically, then this residue pair was saved to the coordination residue pair library. (5) A set-reduction algorithm was used to search for four- or five-residue combinations (containing at least two cysteines) that could coordinate on the same Cd²⁺ based on the coordination residue pair library. The average Cd²⁺ position (larger yellow sphere) was calculated from the individual coordination position for each residue (smaller yellow sphere). (6) The Cd²⁺ position was fine-tuned around the initial coordination, and a scoring function was used to rank the design results.

“perfect” thresholds, then the residue was assigned a higher score, otherwise it was assigned a lower score (Table S1). The summation of scores of all the coordination residues was used for fine-tuning the Cd²⁺-binding position and sorting all the retained solutions.

Using this algorithm, four Cd²⁺ coordination sites with corresponding mutations were identified, which were labeled as CdRBP1 to CdRBP4 (Fig. S1b). All the four Cd²⁺-binding sites were located around the ribose-binding cavity aperture and did not block ribose binding (Fig. S1b). In addition to these mutants, two more mutants were designed with manual inspections in which ribose binding was blocked through introduction of A137Y and R141L into the top two ranked computational search results, CdRBP1 and CdRBP2, respectively. The hydrogen bonding interactions between Arg 141 and ribose were removed by changing to Leucine, and the mutation A137Y result in bump with ribose (Fig. S1c). These mutants were labeled as CdRBP1m and CdRBP2m. All six designs were examined experimentally.

RBP mutation leads to increased Cd²⁺ tolerance. We assessed whether the RBP mutants could increase the toxicity tolerance of engineered *E. coli* strains on the plate by binding Cd²⁺ and reduce its harmful effects on *E. coli* cells. To this end, the RBP mutants were overexpressed into *E. coli*, and then cadmium adsorption of *E. coli* cells expressing wild-type and mutant RBP was measured using a plate screening assay. The growth of *E. coli* expressing wild-type RBP was significantly suppressed when exposed to 150 μM Cd²⁺ (Fig. S2a) due to the toxic effects of the ion. However, *E. coli*

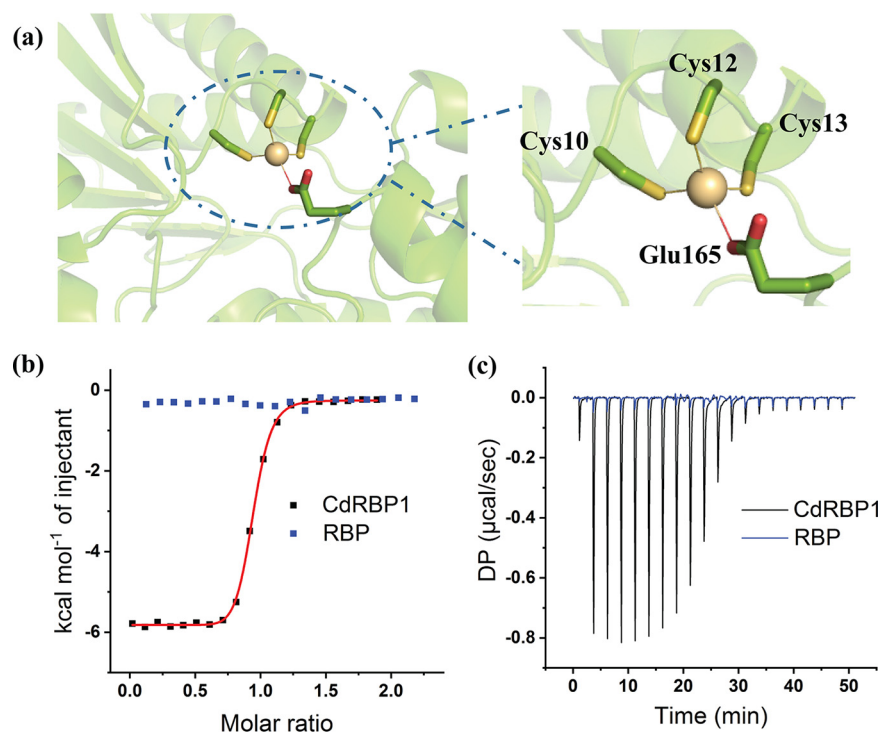


FIG 3 CdRBP1 binding to Cd²⁺. (a) Schematic of the designed Cd²⁺ binding site of CdRBP1 based on the RBP scaffold (left) and residues involved in Cd²⁺ coordination (right). (b) ITC measurements of CdRBP1 (black dots) and wild-type RBP (blue dots) binding with Cd²⁺. The CdRBP1 binding constant was determined by fitting the thermal data from one Cd²⁺ site binding model using MicroCal Origin software. (c) The heat released from each injection during titration of CdRBP1 (black line) and wild-type RBP (blue line).

expressing CdRBP1, CdRBP3, CdRBP1m, and CdRBP2m remained viable under this condition (Fig. S2b), demonstrating that these mutants might be strongly associated with Cd²⁺. We chose these four mutants for further investigation.

Cd²⁺ binding affinity of CdRBPs. Next, we purified and characterized the four CdRBP proteins that showed greatest potential for Cd²⁺ binding. The circular dichroism spectra of these four proteins showed secondary structure features similar to wild-type RBP (Fig. S5). We then performed isothermal titration calorimetry (ITC) experiments to measure the binding affinities and stoichiometric ratios of these proteins with Cd²⁺. The four mutants showed different Cd²⁺ binding behaviors (Fig. 3c and Fig. S3). The ITC titration data demonstrated that CdRBP1, CdRBP1m, and CdRBP2m bound Cd²⁺ in a 1:1 ratio while CdRBP3 exhibited two-site binding behavior (Table 1). The strongest binding protein, CdRBP1, bound Cd²⁺ with a dissociation constant of 268 nM. Under the same experimental conditions, wild-type RBP did not produce any binding signal in the ITC assay (Fig. 3b and 3c). We further verified Cd²⁺ binding of CdRBP1 and CdRBP2m by performing Fourier transform-mass spectrometry analysis. After incubation with Cd²⁺, both proteins showed a 110.2 Dalton increase in molecular weight, indicating complete incorporation of Cd²⁺ into the proteins (Fig. S4a-d). No increase of a Cd²⁺ molecular weight adduct was observed for the purified wild type RBP (Fig. S4e-f).

TABLE 1 Characterization of the Cd²⁺-binding properties of RBP mutants by isothermal titration calorimetry

Variants	Mutant residues	Binding mode	Dissociation constant K_D (M)	Stoichiometric ratio N^a
CdRBP 1	D165E N13C N12C T10C	One site	$(2.68 \pm 0.25) \times 10^{-7}$	1.00 ± 0.00
CdRBP 3	D165E N41H N13C N12C	Two sites	$1 \times 10^{-5}, 1.1 \times 10^{-4}$	ND ^b
CdRBP 1m	D165E N13C N12C T10C A137Y R141L	One site	$(1.00 \pm 0.33) \times 10^{-6}$	1.00 ± 0.01
CdRBP 2m	I132E N41C T19H S9C A137Y R141L	One site	$(2.19 \pm 0.14) \times 10^{-6}$	0.99 ± 0.02

^aN is the apparent stoichiometry from experimental data fitting. \pm , statistical error from fitting.

^bND, not determined. The stoichiometric ratio for CdRBP3 with two binding sites could not be fitted.

Verification of the Cd²⁺ binding sites using alanine scanning mutagenesis. In our computational models, CdRBP1 and CdRBP2m bound to Cd²⁺ through three coordination residues on the N-terminal domain and one glutamate residue on the C-terminal domain (Table 1). To verify the involvement of these residues, we individually replaced them with alanine and then measured the Cd²⁺ binding affinity. The ITC titration experiments indicated that the alanine substitution of Cys10 and Cys13 in CdRBP1 resulted in an approximately 20-fold decrease in binding affinity (Fig. S6). Alanine substitution of Glu165 resulted in a 5-fold decrease. Taken together, these results suggest that Cys10 and Cys13 contribute more to the association energy with Cd²⁺ than Glu165. Alanine substitution of the coordination residue Cys12 resulted in a change in stoichiometric ratio for Cd²⁺ binding, indicating that two types of cadmium binding sites with different binding affinities may present (Fig. S6). We propose that this residue may help orient the flexible loop (residues 8–13) into the proper conformation and promote single-site Cd²⁺ binding. These alanine scanning results suggested that CdRBP1 adopts the computationally designed coordination structure. However, for CdRBP2m, substitution of Glu132 had little effect on Cd²⁺-binding affinity, indicating that the three coordination residues from the N-terminal domain are the primary contributors to Cd²⁺ binding ability (Fig. S6).

Chemotactic response of engineered *E. coli* toward cadmium ions. Since CdRBP1 and CdRBP2m exhibited high Cd²⁺ binding affinity, we tested whether the *E. coli* cells with either of these two mutants expressed show chemotactic response toward Cd²⁺. The responses of the *E. coli* cells were measured using a designed microfluidic chip (Fig. 4a). In these experiments, green fluorescent protein (GFP)-labeled cells were added to a sink hole, and different concentrations of Cd²⁺ were introduced into the source holes (Fig. 4a). Our data demonstrate that the *E. coli* RP437 chemotaxis strain with periplasmic expression of CdRBP1 (RP4371) exhibited higher fluorescence intensity at the end of the observation channel compared with the blank buffer (Fig. 4b), and the fluorescence intensity increase was time dependent (Fig. 4c). These results indicated that RP4371 sensed the Cd²⁺ gradient and responded by accumulating gradually in the high-concentration region. The chemotactic response was also concentration dependent as the fluorescence intensity increased when the Cd²⁺ concentration in the source holes increased from 50 to 300 μ M (Fig. 4d). The wild-type RBP-expressing RP437 strain did not respond to any Cd²⁺ concentration along the gradient (Fig. S7a), thus providing support for the involvement of CdRBP1 in the chemoattractant response toward Cd²⁺. No obvious chemotactic behavior was observed in cells expressing CdRBP2m under the same conditions (Fig. S7b-c).

We further assessed whether these two RBP mutants were exported to the periplasmic region, as this is mandatory for successful chemotactic sensing. We found that the expression level of CdRBP1 was higher than that of CdRBP2m in the periplasmic region of the engineered *E. coli* (Fig. S7d). The reduced CdRBP2m translocation was probably due to the N41C mutation, which was reported to disrupt the periplasmic translocation signal (31).

No chemotactic behaviors were observed for the CdRBP1-expressing *E. coli* VS2811 strain, which lacks the Trg receptor (Δ trg) (Fig. S7e), suggesting that the Trg chemoreceptor is required for the Cd²⁺ response. We also measured the effects of single coordination site mutations of CdRBP1 on *E. coli* chemotactic behavior. We found that all mutants had significantly reduced responses to Cd²⁺ (Fig. S7f-i).

Measurement of domain closure using fluorescence resonance energy transfer assay. To investigate the Cd²⁺-induced conformational changes in CdRBP1, we fused a cyan fluorescent protein (CFP) and a yellow fluorescent protein (YFP) to the N- and C-termini of the protein, respectively. Due to the relative positions of the two domains of RBP, fluorescence resonance energy transfer (FRET) occurs when the RBP structure is open (the apo state without ligand binding) (42, 43). Upon Cd²⁺ binding, domain closure occurs, which drives the two fluorescent proteins away from each other, resulting in a decreased FRET signal (Fig. 5a),

As shown in Fig. 5b, CdRBP1 exhibited a lower FRET ratio upon Cd²⁺ binding that was similar to RBP upon ribose binding, implying that CdRBP1 and the wild-type protein undergo similar conformational changes.

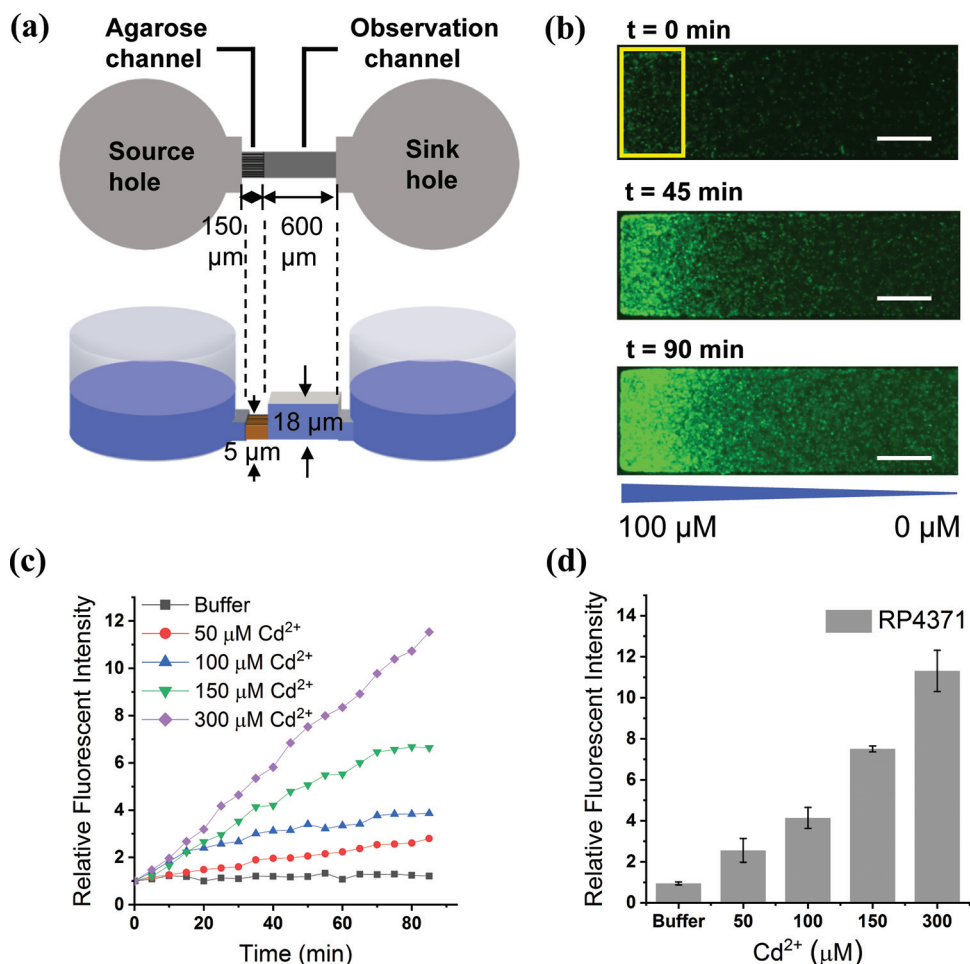


FIG 4 Chemotaxis of a CdRBP1-expressing *E. coli* strain toward Cd²⁺. (a) Schematic of the microfluidic device used to examine chemotaxis (side view and top view). (b) Responses of *E. coli* RP4371 to a Cd²⁺ gradient, recorded using fluorescence microscopy imaging at different times (scale bar, 100 μm). The Cd²⁺ gradient was from 0 μM to 100 μM, run through the observation channel. (c) Responses of the RP4371 strain to different Cd²⁺ concentrations as a function of time. The response is represented by relative fluorescent intensity, compared with $t = 0$ min at the same Cd²⁺ concentration in the analysis region (yellow rectangle, scale, 100 μm × 200 μm). The indicated concentrations are the source concentrations added to the source holes. (d) Final relative fluorescent intensities in the presence of different Cd²⁺ concentrations (mean ± SEM; N = 3).

Metal ion selectivity of the engineered RP4371 strain. We further tested the metal ion selectivity of our engineered Cd²⁺-sensing RP4371 strain. Among the 10 commonly encountered metal ions, including Co²⁺, Fe³⁺, Ca²⁺, Mn²⁺, Mg²⁺, Ni²⁺, Cu²⁺, Fe²⁺, and Zn²⁺, RP4371 only significantly responds to Cd²⁺ in the microfluidic chemotaxis assay (Fig. 6). The great selectivity for Cd²⁺ of the engineered *E. coli* strain might result from the binding specificity of CdRBP1 to Cd²⁺ and the correct conformational changes of CdRBP1 caused by the Cd²⁺ binding. We used ITC to measure the *in vitro* binding affinity of CdRBP1 to metal ions with similar coordination properties to Cd²⁺. Compared to its affinity to Cd²⁺, CdRBP1 showed a much weaker binding affinity to Ni²⁺ (approximately 10-fold lower) and no binding to Mg²⁺ (Fig. S7j). For the most similar ion, Zn²⁺, our ITC study showed that it binds CdRBP1 with exothermic profile followed by endothermic profile with increasing zinc ion concentration (Fig. S7l), indicating that there are two types of zinc binding sites in CdRBP1. The first binding site, which is the stronger binding site, was fitted, and the dissociation constant was 5.68×10^{-7} M, which is about 2-fold weaker than that of Cd²⁺, as shown in Fig. S7j. FRET assay demonstrated that Zn²⁺ binding did not produce similar conformational change in CdRBP1 (Fig. S7m), which is consistent with the chemotaxis result.

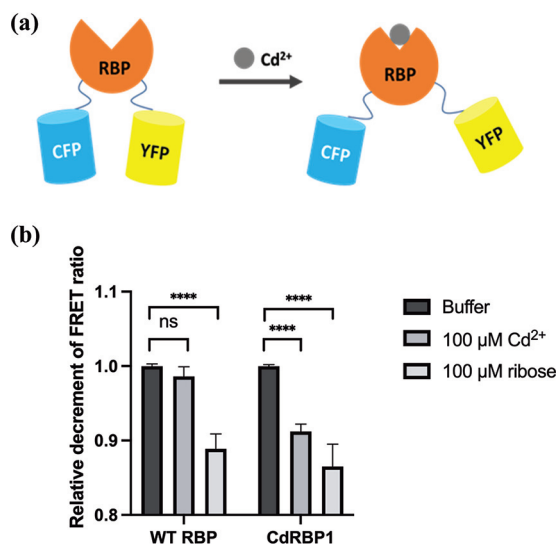


FIG 5 Interdomain conformational changes were measured using a FRET assay. (a) Schematic illustration of the FRET assay based on RBP structure. (b) Changes in FRET ratio of purified wild-type RBP and CdRBP1 in the presence of ribose or Cd²⁺ compared with proteins in buffer alone (mean \pm SEM; N = 5; ****, $P < 0.0001$, unpaired two-tailed t test; ns, not significant).

DISCUSSION

Bacterial chemotactic intelligence provides a natural system for reconstructing sensory robots that autonomously hunt for novel ligands. Here, we present a strategy for designing and developing Cd²⁺-binding proteins based on the *E. coli* RBP, a periplasmic binding protein responsible for chemotaxis. By employing a computational design method to search for Cd²⁺ binding sites that accommodate proper coordination geometry, we obtained four Cd²⁺-binding RBP variants. Using a FRET assay, we confirmed that in one of the computationally designed proteins, namely, CdRBP1, Cd²⁺ binding induced domain closure reproducing the conformational change of the wild-type RBP with ribose-binding. This combination of computational and experimental methods represents a successful example of incorporation of Cd²⁺ into the RBP scaffold.

Many attempts have been made to construct RBP mutants that bind to novel ligands at the ribose-binding site. However, due to the tight binding of ribose and the complicated conformational changes that occur after ribose binding, some mutations on RBP result in misfolded or poorly translocated proteins (31), making these designs produce undesired binding modes or binding abilities (17, 18). Recently, RBPs were engineered to

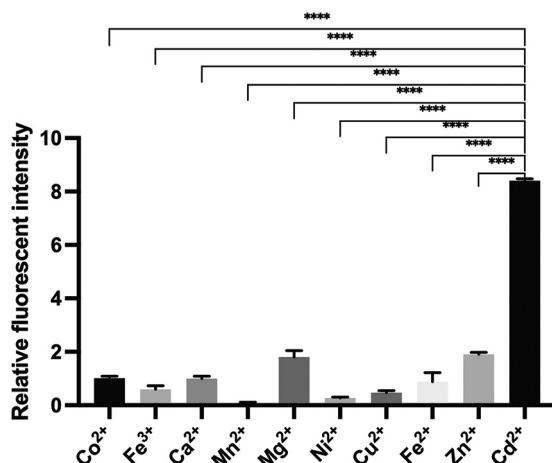


FIG 6 CdRBP1 binding is selective for Cd²⁺. A microfluidic chemotaxis assay was performed to examine the response of *E. coli* RP4371 to 250 μM of the indicated metal ion (mean \pm SEM; N = 3; ****, $P < 0.0001$, unpaired two-tailed t test).

bind compounds that are structurally similar to ribose at the ribose binding site, by mutant library screening strategy (44). The resulting RBP variants were shown to bind with ligand *in vitro*. Unfortunately, the binding strength is much weaker and the effect on chemotaxis signal transduction pathway is unclear. These previous efforts indicate that engineering RBP to bind new ligands within the ribose-binding site and further related to function remains challenging. Instead of modifying the ribose binding site, we designed a Cd^{2+} -binding site in RBP at a new site, that is, the lids of the domain closure structure. Our mutation studies showed that the right coordination of Cd^{2+} at this site is important for binding and to induce the necessary domain closure conformational changes required by the chemotaxis function. For example, when we tried to improve Cd^{2+} binding affinity of CdRBP1 by mutating D165 to Cys and elongating the nearby loop to make the sulfur atom reaching Cd^{2+} (Fig. S5), the mutants tend to bind two Cd^{2+} ions (Fig. S5d). The mutation of 10C on CdRBP1 to 41H (resulting CdRBP3) also changed the binding stoichiometry to two Cd^{2+} ions with one protein (Table 1, Fig. S3). Alanine substitution of each coordination residue on CdRBP1 resulted in either decrease of binding affinity or binding mode changes (Fig. S6). These mutants showed significantly reduced chemotactic responses to Cd^{2+} in cells (Fig. S7). The CdRBP2m variant possesses favorable binding affinity with Cd^{2+} *in vitro* but showed no chemotactic sensing in cells (Fig. S7b). One possible reason is the disrupted periplasmic translocation ability. In addition, another plausible reason is due to lack of interaction with the Trg receptor. We checked the binding of CdRBP2m with ribose using FRET assay. As shown in Fig. S7k, CdRBP2m exhibited less decrement of FRET ratio upon ribose binding compared to wild-type binding with ribose, implying that mutations in CdRBP2m impair ribose binding-induced conformational change. When CdRBP2m binds Cd^{2+} , an opposite change in the FRET ratio was observed, indicating a more open conformation that is distinct from the closed conformation required by Trg binding. Thus, Cd^{2+} binding of CdRBP2m does not produce necessary conformational change required for chemoreceptor binding; this should be the main reason for its lack of chemotaxis activity (31). Taken together, besides tight binding, proper conformational change and the right translocation to the periplasmic region of variants are mandatory for the successful chemotactic sensing.

Chemotaxis signaling produced by RBP requires its conformational change from open conformation to close conformation. Thus, only strong binding does not necessarily result in chemotaxis signaling. The large conformational change itself produces entropy loss during binding. In addition, the interface between the two domains of RBP are dominantly composed of polar residues, which are more solvent exposed on protein opening status (27). The change of protein conformation induced by cadmium would squeeze out water molecules. The penalty of desolvation of polar groups next to the protein-solvent interface upon binding would give rise to the decrease of observed binding affinity. Natural cadmium binding proteins, such as metallothionein, have rigid structures that do not change after cadmium binding. These proteins usually use multiple thiols from cysteine residues to bind Cd^{2+} , forming Cd^{2+} thiolate clusters with high binding strength. Thus, both the large conformational change of RBP and the burying of polar groups upon cadmium binding produce entropic loss that result in weakened binding compared to natural cadmium binding proteins that use rigid binding sites.

Using our in-house-designed microfluidic device for the bacterial chemotaxis assay, we found that engineered *E. coli* expressing CdRBP1 exhibited chemotactic behavior along a Cd^{2+} gradient with divalent metal ion selectivity. Our approach provides a strategy for designing novel bacterial chemotaxis properties based on PBPs. Heavy metals are chemorepellents for most natural microbes; our work demonstrates that it is possible to engineer bacteria with chemoattractant properties toward metal ions (Cd^{2+}). With the benefit of the bacteria moving toward Cd^{2+} , our engineered *E. coli* systems may improve the current Cd^{2+} microbial sensors and remediation strategies, which are vulnerable due to the toxic nature of Cd^{2+} (25). In increased levels of contaminated environment, durable cadmium detection and remediation might be realized by incorporation of proteins with high cadmium-binding-affinity or cadmium detection circuits into our engineered *E. coli* cells.

MATERIALS AND METHODS

Computational design of Cd²⁺ binding proteins. The structure of the RBP was divided into two domains, domain I (residues 1–103, 236–263) and domain II (residues 104–235, 264–271). Only residues on the domain–domain interface were mutable for cadmium coordination. Prolines and residues in the hinge region (102–104, 234–236, 263–265), which are important for maintaining proper conformation, were set to be nonmutable. If a C α atom in a domain was within 12 Å of any C α atom of the other domain, it was defined as interfacial. A backbone-dependent rotamer library from Dunbrack's lab was used for side chain assembly (36). The T132E, S9C, T10H, and N41C mutations are shown in Fig. 2. Before side chain assembly, backbone conformations were slightly perturbed using the backrub method for considering backbone flexibility (45), and the backrub angle was set as -6° , 0° , or $+6^\circ$.

The geometry parameters in Table S1 were used to calculate the standard coordination sites for a single residue. Residues with cadmium coordination sites close to each other (within 1.6 Å) were combined (as Fig. 2 step 5 illustrates) and screened with the “acceptable” thresholds for each geometry parameter (bond, angle, or dihedral) in Table S1. A set-reduction algorithm was used for the combinations search based on the coordination residue pair library. There is a library reduction process on the initial coordination residue pair library. Using the initial coordination pair library, the maximum number of possible co-ordination residues for a given residue in the coordination library was calculated. If the maximum co-ordination number of a residue was less than 3, it was removed from the library. After reducing the coordination residue library, the coordination residue pair library was updated.

Solutions meeting the “acceptable” thresholds were scored by adding up the coordination score of each residue. The coordination scoring function expresses the quality of residue conformation and coordination geometry as

$$\text{Score} = \sum_i \text{Quality}_i + \text{CoordinationScore}_i \quad (1)$$

The quality of residue conformation includes side chain conformation, backbone conformation, and coordination type terms as

$$\text{Quality}_i = 0.9 \times \log(pi) - 2 \times \frac{|\theta_{1,3}|}{6.0} + w_i \quad (2)$$

where p is the probability of the side chain rotamer and $\theta_{1,3}$ is the backbone backrub angle. The coordination type weight, w , was defined as Cys, 0.8; His, 1.0; Asp/Glu, 0.8; and backbone oxygen 3.0.

The coordination geometry was scored as a two-level function, namely, perfect and acceptable. There was a threshold for all types of coordination at these two levels (see Table S1). For perfect and acceptable coordination, a discrete score (see Table S1 for the different types of coordination) was added to the total score. The location of the cadmium site was fine-tuned around the initial coordination in a $0.8 \times 0.8 \times 0.8$ Å lattice within a 0.1 Å space.

Plasmid construction for RBP and variants. The genes encoding the RBP and RBP variants were cloned into expression vectors by PCR followed by restriction enzyme digestion and ligation. All RBP variants were obtained using a site-directed mutagenesis kit (TIANGEN). All sequences were confirmed by DNA sequencing. See the SI for experimental details.

Plate sensitivity assay. The number of *E. coli* strains overexpressing wild type RBP and variants were normalized and subjected to six 10-fold serial dilutions. Samples (10 μ L) of each were spotted onto LB agar plates containing different cadmium concentrations. All plates were incubated at 37°C for 16 h before measurement. See the supplemental material for experimental details.

Protein expression and purification. The recombinant plasmids expressing wild-type RBP, cadmium-binding proteins, or the single alanine substitution mutants in pET-28a were transformed individually into BL21(DE3) strains for protein expression. Cell lysate was applied to HisTrap columns (GE Healthcare Life Sciences) and the N-terminal 6 \times His-tag was removed by on-column digestion. The target protein was eluted and then applied to a Sephacryl S-200 HR column for further purification. Proteins with FRET pairs were expressed and purified using a similar method.

Circular dichroism spectroscopy. The far-UV CD spectra (190–260 nm) were recorded using MOS 450 AF/CD (Biologic, France) at room temperature. Protein samples were diluted to a final concentration of 0.1 mg/mL using 20 mM phosphate potassium buffer (pH 7.3) containing 150 mM KCl. Samples were scanned in a quartz cuvette with the pathlength of 1 mm at a spectral bandwidth of 1 nm. Signals from three scans were averaged.

Isothermal titration calorimetry (ITC) assay. An iTC200 Microcalorimeter (MicroCal, USA) was used to measure the binding affinity (K_a) of cadmium with purified wild-type RBP and the designed variants and single alanine mutants. All data were analyzed using MicroCal Origin software. See the supplemental material for details.

FT-mass spectrometry analysis. The total mass of cadmium-free proteins and cadmium-binding proteins was determined by using a Fourier Transform Ion Cyclotron Resonance mass spectrometer with electrospray ionization (Solarix XR, Bruker). Cadmium-binding proteins were prepared by incubating the proteins with CdCl₂ for 30 min at 4°C before desalting. All proteins were desalted by ultrafiltration with ddH₂O three times before measurement.

FRET assay. The FRET protein samples were diluted to 5 μ M and then incubated with excessive cadmium ions or ribose to a final concentration of 100 μ M at room temperature for 20 min. ECFP was excited at 433 nm, and then the emission fluorescence signal intensity was measured at 475 nm and

527 nm for ECFP and EYFP, respectively. Fluorescence signals were recorded on a microplate reader (Synergy, Biotek). The FRET ratio was calculated by dividing the fluorescent emission intensity of ECFP by that of EYFP. Proteins diluted in blank buffer were measured as a control.

Microfluidic experiments. Details of the microfluidic device and operating principles are shown in Fig. 5a and the supplemental material. *E. coli* cells expressing GFP protein were loaded into the sink holes and then allowed to diffuse into the observation channel to reach a steady state. Cadmium and other divalent metal ions were added into the source holes. The responses of the cells to different concentrations of metal ions were observed by microscopy for 1 to 2 h. The response of the cells was quantified by measuring the fluorescent intensity of the analysis region using the NIH Image J software (46).

Analysis of protein expression in periplasmic region by osmotic shock. Recombinant plasmids expressing RBP and variants with N-terminal signal peptide sequence were transformed into *E. coli* RP437 and induced with sodium salicylate. The proteins expressed in the periplasmic region were extracted using an osmotic shock procedure (46, 47). The extracted protein samples were then analyzed by 10% SDS-polyacrylamide gel electrophoresis (PAGE). See the supplemental material for experimental details.

Strains and plasmids. Information regarding the genotypes, phenotypes, and sources of the bacterial strains and plasmids used in this study are listed in Table S2.

Data availability. The source code of the design program and the designed structure models are available at <https://github.com/victorPKU/CADEIN>.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TEXT S1, DOCX file, 0.03 MB.

FIG S1, TIF file, 2 MB.

FIG S2, TIF file, 1.8 MB.

FIG S3, TIF file, 2.7 MB.

FIG S4, TIF file, 1.5 MB.

FIG S5, TIF file, 2.6 MB.

FIG S6, TIF file, 2.6 MB.

FIG S7, TIF file, 1.1 MB.

TABLE S1, DOCX file, 0.1 MB.

TABLE S2, DOCX file, 0.01 MB.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (21633001). We thank Professor J. S. Parkinson (University of Utah) for the RP437 *E. coli* strain and Professor Victor Sourjik (Max Planck Institute for Terrestrial Microbiology) for the VS281 *E. coli* strain and the PSN77 plasmid used in this study. We thank Dr. Shuangyu Bi (Max Planck Institute for Terrestrial Microbiology) for helpful discussions and suggestions. We also thank Professor Chunxiang Luo (Peking University) for assistance with the microfluidic experiments.

L. Lai conceived the project and supervised the work. H. Li designed and performed the experiments and data analysis. C. Zhang performed the computational work. X. Chen designed the microfluidic chip. H. You participated in the plate sensitivity assay and FT-MS experiment. H. Li, C. Zhang, and L. Lai wrote the manuscript with inputs from all authors.

We declare that there are no conflicts of interest regarding the publication of this article.

REFERENCES

- Wadhams GH, Armitage JP. 2004. Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol* 5:1024–1037. <https://doi.org/10.1038/nrm1524>.
- Stock JB, Surette MG, McCleary WR, Stock AM. 1992. Signal transduction in bacterial chemotaxis. *J Biol Chem* 267:19753–19756. [https://doi.org/10.1016/S0021-9258\(19\)88615-5](https://doi.org/10.1016/S0021-9258(19)88615-5).
- Typas A, Sourjik V. 2015. Bacterial protein networks: properties and functions. *Nat Rev Microbiol* 13:559–572. <https://doi.org/10.1038/nrmicro3508>.
- Jacob-Dubuisson F, Mechaly A, Betton JM, Antoine R. 2018. Structural insights into the signalling mechanisms of two-component systems. *Nat Rev Microbiol* 16:585–593. <https://doi.org/10.1038/s41579-018-0055-7>.
- Bren A, Eisenbach M. 2000. How signals are heard during bacterial chemotaxis: protein-protein interactions in sensory signal propagation. *J Bacteriol* 182:6865–6873. <https://doi.org/10.1128/JB.182.24.6865-6873.2000>.
- Ahmad F, Zhu DC, Sun JZ. 2021. Environmental fate of tetracycline antibiotics: degradation pathway mechanisms, challenges, and perspectives. *Environ Sci Eur* 33:64. <https://doi.org/10.1186/s12302-021-00505-y>.
- Pandey G, Jain RK. 2002. Bacterial chemotaxis toward environmental pollutants: role in bioremediation. *Appl Environ Microbiol* 68:5789–5795. <https://doi.org/10.1128/AEM.68.12.5789-5795.2002>.
- Shao J, Xuan M, Zhang H, Lin X, Wu Z, He Q. 2017. Chemotaxis-guided hybrid neutrophil micromotors for targeted drug transport. *Angew Chem Int Ed Engl* 56:12935–12939. <https://doi.org/10.1002/anie.201706570>.
- Bi S, Pollard AM, Yang Y, Jin F, Sourjik V. 2016. Engineering hybrid chemotaxis receptors in bacteria. *ACS Synth Biol* 5:989–1001. <https://doi.org/10.1021/acssynbio.6b00053>.
- Bi S, Yu D, Si G, Luo C, Li T, Ouyang Q, Jakovljevic V, Sourjik V, Tu Y, Lai L. 2013. Discovery of novel chemoeffectors and rational design of *Escherichia coli* chemoreceptor specificity. *Proc Natl Acad Sci U S A* 110:16814–16819. <https://doi.org/10.1073/pnas.1306811110>.

11. Dwyer MA, Hellinga HW. 2004. Periplasmic binding proteins: a versatile superfamily for protein engineering. *Curr Opin Struct Biol* 14:495–504. <https://doi.org/10.1016/j.sbi.2004.07.004>.
12. Scheepers GH, Lycklama ANJA, Poolman B. 2016. An updated structural classification of substrate-binding proteins. *FEBS Lett* 590:4393–4401. <https://doi.org/10.1002/1873-3468.12445>.
13. Quioco FA, Ledvina PS. 1996. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol Microbiol* 20:17–25. <https://doi.org/10.1111/j.1365-2958.1996.tb02484.x>.
14. Gouridis G, Schuurman-Wolters GK, Ploetz E, Husada F, Vietrov R, de Boer M, Cordes T, Poolman B. 2015. Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nat Struct Mol Biol* 22:57–64. <https://doi.org/10.1038/nsmb.2929>.
15. Shilton BH, Flocco MM, Nilsson M, Mowbray SL. 1996. Conformational changes of three periplasmic receptors for bacterial chemotaxis and transport: the maltose-, glucose/galactose- and ribose-binding proteins. *J Mol Biol* 264:350–363. <https://doi.org/10.1006/jmbi.1996.0645>.
16. Borrok MJ, Zhu YM, Forest KT, Kiessling LL. 2009. Structure-based design of a periplasmic binding protein antagonist that prevents domain closure. *ACS Chem Biol* 4:447–456. <https://doi.org/10.1021/cb900021q>.
17. Reimer A, Yagur-Kroll S, Belkin S, Roy S, van der Meer JR. 2014. *Escherichia coli* ribose binding protein based bioreporters revisited. *Sci Rep* 4:9. <https://doi.org/10.1038/srep05626>.
18. Schreier B, Stumpp C, Wiesner S, Hocker B. 2009. Computational design of ligand binding is not a solved problem. *Proc Natl Acad Sci U S A* 106:18491–18496. <https://doi.org/10.1073/pnas.0907950106>.
19. Peng X, Du J, Fan J, Wang J, Wu Y, Zhao J, Sun S, Xu T. 2007. A selective fluorescent sensor for imaging Cd²⁺ in living cells. *J Am Chem Soc* 129:1500–1501. <https://doi.org/10.1021/ja0643319>.
20. Liu Z, Zhang C, He W, Yang Z, Gao X, Guo Z. 2010. A highly sensitive ratiometric fluorescent probe for Cd²⁺ detection in aqueous solution and living cells. *Chem Commun (Camb)* 46:6138–6140. <https://doi.org/10.1039/c0cc00662a>.
21. Vinkenborg JL, van Duijnhoven SM, Merckx M. 2011. Reengineering of a fluorescent zinc sensor protein yields the first genetically encoded cadmium probe. *Chem Commun (Camb)* 47:11879–11881. <https://doi.org/10.1039/c1cc14944j>.
22. Guo Y, Zhang Y, Shao H, Wang Z, Wang X, Jiang X. 2014. Label-free colorimetric detection of cadmium ions in rice samples using gold nanoparticles. *Anal Chem* 86:8530–8534. <https://doi.org/10.1021/ac502461r>.
23. Liu Q, Yuan F, Liang Y, Li Z. 2015. Cadmium adsorption by *E. coli* with surface displayed CadR. *RSC Adv* 5:16089–16092. <https://doi.org/10.1039/C4RA07805E>.
24. Wu Q, Zhou M, Shi J, Li Q, Yang M, Zhang Z. 2017. Synthesis of water-soluble Ag₂S quantum dots with fluorescence in the second near-infrared window for turn-on detection of Zn(II) and Cd(II). *Anal Chem* 89:6616–6623. <https://doi.org/10.1021/acs.analchem.7b00777>.
25. Bereza-Malcolm LT, Mann G, Franks AE. 2015. Environmental sensing of heavy metals through whole cell microbial biosensors: a synthetic biology approach. *ACS Synth Biol* 4:535–546. <https://doi.org/10.1021/sb500286r>.
26. Singh R, Olson MS. 2008. Application of bacterial swimming and chemotaxis for enhanced bioremediation. *In* Shah V (ed), *Emerging environmental technologies*. Springer, Dordrecht, the Netherlands. https://doi.org/10.1007/978-1-4020-8786-8_7.
27. Bjorkman AJ, Mowbray SL. 1998. Multiple open forms of ribose-binding protein trace the path of its conformational change. *J Mol Biol* 279:651–664. <https://doi.org/10.1006/jmbi.1998.1785>.
28. Grebe TW, Stock J. 1998. Bacterial chemotaxis: the five sensors of a bacterium. *Curr Biol* 8:R154–R157. [https://doi.org/10.1016/S0960-9822\(98\)00098-0](https://doi.org/10.1016/S0960-9822(98)00098-0).
29. Binnie RA, Zhang H, Mowbray S, Hermodson MA. 1992. Functional mapping of the surface of *Escherichia coli* ribose-binding protein: mutations that affect chemotaxis and transport. *Protein Sci* 1:1642–1651. <https://doi.org/10.1002/pro.5560011212>.
30. Vercillo NC, Herald KJ, Fox JM, Der BS, Dattelbaum JD. 2007. Analysis of ligand binding to a ribose biosensor using site-directed mutagenesis and fluorescence spectroscopy. *Protein Sci* 16:362–368. <https://doi.org/10.1110/ps.06259570.7>.
31. Reimer A, Maffenbeier V, Dubey M, Sentchilo V, Tavares D, Gil MH, Beggah S, van der Meer JR. 2017. Complete alanine scanning of the *Escherichia coli* RbsB ribose binding protein reveals residues important for chemoreceptor signaling and periplasmic abundance. *Sci Rep* 7:7. <https://doi.org/10.1038/s41598-017-08035-5>.
32. Zhou L, Bosscher M, Zhang CS, Ozcubukcu S, Zhang L, Zhang W, Li CJ, Liu JZ, Jensen MP, Lai LH, He C. 2014. A protein engineered to bind uranyl selectively and with femtomolar affinity. *Nat Chem* 6:236–241. <https://doi.org/10.1038/nchem.1856>.
33. Brint AT, Willett P. 1987. Pharmacophoric pattern matching in files of 3D chemical structures: comparison of geometric searching algorithms. *J Mol Graph* 5:49–56. [https://doi.org/10.1016/0263-7855\(87\)80045-0](https://doi.org/10.1016/0263-7855(87)80045-0).
34. Friedman R. 2014. Structural and computational insights into the versatility of cadmium binding to proteins. *Dalton Trans* 43:2878–2887. <https://doi.org/10.1039/c3dt52810c>.
35. Bjorkman AJ, Binnie RA, Zhang H, Cole LB, Hermodson MA, Mowbray SL. 1994. Probing protein-protein interactions: the ribose-binding protein in bacterial transport and chemotaxis. *J Biol Chem* 269:30206–30211. [https://doi.org/10.1016/S0021-9258\(18\)43798-2](https://doi.org/10.1016/S0021-9258(18)43798-2).
36. Shapovalov MV, Dunbrack RL, Jr. 2011. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* 19:844–858. <https://doi.org/10.1016/j.str.2011.03.019>.
37. Cuyppers MG, Mason SA, Mossou E, Haertlein M, Forsyth VT, Mitchell EP. 2016. Macromolecular structure phasing by neutron anomalous diffraction. *Sci Rep* 6:31487. <https://doi.org/10.1038/srep31487>.
38. Maher M, Cross M, Wilce MC, Guss JM, Wedd AG. 2004. Metal-substituted derivatives of the rubredoxin from *Clostridium pasteurianum*. *Acta Crystallogr D Biol Crystallogr* 60:298–303. <https://doi.org/10.1107/S090744490302794X>.
39. Hingorani K, Pace R, Whitney S, Murray JW, Smith P, Cheah MH, Wydrzynski T, Hillier W. 2014. Photo-oxidation of tyrosine in a bio-engineered bacterioferritin 'reaction centre'—a protein model for artificial photosynthesis. *Biochim Biophys Acta* 1837:1821–1834. <https://doi.org/10.1016/j.bbabi.2014.07.019>.
40. Kaus-Drobek M, Czapinska H, Sokolowska M, Tamulaitis G, Szczepanowski RH, Urbanek C, Siksnys V, Bochtler M. 2007. Restriction endonuclease MvaI is a monomer that recognizes its target sequence asymmetrically. *Nucleic Acids Res* 35:2035–2046. <https://doi.org/10.1093/nar/gkm064>.
41. Tarbouriech N, Prates JA, Fontes CM, Davies GJ. 2005. Molecular determinants of substrate specificity in the feruloyl esterase module of xylanase 10B from *Clostridium thermocellum*. *Acta Crystallogr D Biol Crystallogr* 61:194–197. <https://doi.org/10.1107/S0907444904029695>.
42. Lager I, Fehr M, Frommer WB, Lalonde S. 2003. Development of a fluorescent nanosensor for ribose. *FEBS Lett* 553:85–89. [https://doi.org/10.1016/S0014-5793\(03\)00976-1](https://doi.org/10.1016/S0014-5793(03)00976-1).
43. Marshall KE, Robinson EW, Hengel SM, Pasa-Tolic L, Roesijadi G. 2012. FRET imaging of diatoms expressing a biosilica-localized ribose sensor. *PLoS One* 7:e33771. <https://doi.org/10.1371/journal.pone.0033771>.
44. Tavares D, Reimer A, Roy S, Joublin A, Sentchilo V, van der Meer JR. 2019. Computational redesign of the *Escherichia coli* ribose-binding protein ligand binding pocket for 1,3-cyclohexanediol and cyclohexanol. *Sci Rep* 9:16940. <https://doi.org/10.1038/s41598-019-53507-5>.
45. Georgiev I, Keedy D, Richardson JS, Richardson DC, Donald BR. 2008. Algorithm for backrub motions in protein design. *Bioinform* 24:i196–i204. <https://doi.org/10.1093/bioinformatics/btn169>.
46. Rasband WS. 1997–2015. ImageJ. National Institutes of Health, Bethesda, Maryland, USA. <http://imagej.nih.gov/ij>.
47. Nossal NG, Heppel LA. 1966. The release of enzymes by osmotic shock from *Escherichia coli* in exponential phase. *J Biol Chem* 241:3055–3062. [https://doi.org/10.1016/S0021-9258\(18\)96497-5](https://doi.org/10.1016/S0021-9258(18)96497-5).