Engineering supramolecular forming proteins to chelate heavy metals for waste water remediation

[1,2]George L. Sun, [1,2,3]Angela M. Belcher*

[1]Department of Biological Engineering, [2]Koch Institute of Integrative Cancer Research, [3]Department of Material Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Corresponding author: *belcher@mit.edu

Abstract

Physicochemical technologies have dominated water treatment methods, but their impact on reducing contaminated waters has lagged behind the rate of heavy metal waste production. Ion-exchange is one such method that is routinely used industrially but remains difficult to deploy in developing nations. Ion-exchange requires sophisticated design strategies and advanced chemical synthesis techniques which creates a high cost and technical barrier for adoption. To circumvent these limitations, a biological mimic of ion-exchange is investigated, as the mechanism of reversible metal chelation is already ubiquitous and highly evolved in a variety of proteins. Herein, this work shows that proteins from the CTP synthase and glutamine synthetase family, recently discovered to have polymerization properties, can act as metal chelating agents for water purification applications. Specifically, proteins pyrG and glnA show polymerization behavior that aggregate into solid matrixes when exposed to metals. Their surfaces can be decorated with

metal binding domains to tailor specific removal of copper, cadmium, mercury, lead, and even calcium. These proteins can be further modified to harbor binding tags to allow anchorage onto denser substrates to increase metal-protein sedimentation rates. With this system, modified pyrG and glnA show more than 50% removal of 1 millimolar amounts of metal in one round (a quantity more than 100 times higher than EPA standards) and more than 80% recovery of metals. These results show that biological systems can be just as efficacious and more versatile than traditional ion-exchange technologies.

Introduction

Physicochemical technologies are the methods of choice for wastewater remediation, with ion-exchange being one of the most commonly used. Ion-exchange is a well-established technology with high removal efficiency and fast removal kinetics[1–4]. Mechanistically, ion-exchange utilizes resins which are functionalized with multiple copies of metal binding groups (e.g. sulfonic, -SO_3H, and carboxylic acids, -COOH)[1] to chelate metals out of solution. What limits the wide-spread adoption of ion-exchange for global water treatment is the cost, handling, and production of ion-exchange resins[2,5]. Resins are designed synthetically and require rounds of chemical synthesis for testing and production. More so, the design of exchangers is non-trivial, as custom software such as Dow Water's Computer Assisted Design for Ion Exchange (CADIX), Purolite Pure DesignTM, and other programs are sometimes necessary to aid in the design strategy process[6]. What this design flow suggest is the creation, testing, and manufacture of resins is arduous requiring large capital and technical skill. This process also produces secondary waste, as many of the chemical processes are prone to generating wasteful by-products or effluent[5,6].

This work aims to improve the underlying mechanism of ion-exchange, which entails a deeper study of its two components: the resin and the metal binding group. The resin provides the physical medium in which exchangers can be packed and stored as a solid unit, while the metal binder performs the metal chelation and water purification. Both are key and mutually dependent. Without the resin, chelated metals would remain free-floating. Without the chelators, resins have no metal binding property, thus serving no remediation function. Rather than creating this behavior synthetically, everyday biological systems have already evolved sophisticated mechanisms for handling environmental metal content. One such example is the removal of metals through chelation by polymerized chains of phytochelatins, or binding onto de-toxifying metallothionein proteins[7,8]. In addition, the functional groups of strong and weak exchangers are ubiquitous, and often modifiable in most peptides/proteins (e.g. $-S^{2-}$, $-NH_2$, and $-COOH^-$). In previous work, this approach has been attempted by functionalizing cell surfaces with metal binding peptides and proteins. The cells act as the physical anchor (i.e. resin) while the peptides serve as the exchanger. One of the first attempts was in 1996 by Sousa et al. who used bacteria display to anchor poly-His peptides to accumulate metals on the cell surface[9]. Thereon, other works showed similar results by functionalizing bacterial or yeast surfaces with metal-specific proteins such as mercuric reductase or metallothioenins[10–13]. Compared to synthetic resins, biology provides an easier platform to produce resins (i.e. cells) and production and engineerability of exchanges (i.e. peptide/proteins). Given that almost a third of proteins harbor some metal binding pocket[14], the diversity and richness in discovering unique protein-metal chelators is enticing, and the advances in genetic and protein engineering make it possible to further augment protein-metal binding interactions for controlled metal removal applications.

However, past results in biologically mimicked ion-exchange has not yet been able to compete with synthetic ion-exchange with respects to metal removal capacities. Results for cell surface display have capture capacities in the tens of nanomolar of metal per gram of cell dry weight[9,11,13], whereas ion-exchange typical has capture capacities in the millimolar of metal per gram dry weight, almost 3-6 orders of magnitude more[15,16]. As metal removal boils down to a numbers game, the limitation in biologically mimicked ion-exchange is the low number of peptide/protein expression per cell per volume. To illustrate, a typical number of protein expression for bacterial and yeast display range from 1 to 100 thousand[17,18]. Given modest estimates of bacterial or yeast culture density (assuming 1 OD_{600} per mL), and an upper bound of metal uptake assuming each protein has 1 completely occupied binding site, equates to roughly 2-200 nanomolar of metal removal capacity, an almost insignificant amount when dealing with micro to millimolar levels of metal concentrations.

Therefore, to capitalize on the benefits of using biology, e.g. environmentally friendly, diversity, engineerability, etc. this work instead focuses on using proteins discovered to have polymerization properties for heavy metal chelation and sedimentation. Using proteins that polymerize increases protein quantity per volume, and relying on proteins rather than an entire cell for metal capture increase capture capacities per dry weight. Furthermore, these proteins can be fused with metal binding domains to enhance metal capture capacities using modern genetic and protein engineering techniques. When polymerized, these supramolecular complexes decorated with multiple copies of metal binding domains pull down captured metal much like a sponge. The two proteins investigated in this study are bacterial pyrG from the CTP synthase family, and glnA from the glutamine synthetase (GS) family. The existence of these proteins

have been known for decades, however it was not until recently that CTP was observed to form cytoplasmic filaments in a variety of organisms[19–23]. This peculiar behavior has been associated with subcellular protein organization for enzyme regulation[20,22,23]. A multitude of proteins have been screened and discovered to form similar large cytoplasmic supramolecular complexes in *E. coli*, *S. cerevisiae*, and *D. melanogaster*[23–25]. In these screens another protein, GS, was determined to also form supramolecular complexes, and surprisingly this property has been known since the 1970s[25–28]. In the context of metal removal, the particular enzymatic function, and even the cellular role of these proteins, is irrelevant; the main interest is the fact that both CTP and GS have controllable means to initiate polymerization and metal binding, and the focus of this work is aimed to capitalize on this mechanism for a direct application in heavy metal waste removal.

Results

Using CTP and GS for toxic metal removal

Representative crystal structures for pyrG (accession number #P0A7E5) and glnA (#P0A9C5) (PDB ID# 1S1M and 1FPY, respectively) reveal that these proteins are monomers of a larger symmetric structure (**Figure 1a,b**). Past studies have shown that pyrG form monomers, dimers, and tetramers[21] whereas glnA often forms 2 hexagonal structures stacked on top of one another to form a 12-mer structure[27]. Genes for pyrG and glnA were isolated from the genome of a DH5α strain and inserted into IPTG inducible pET28c(+) vector. Confirmed sequences were transformed into BL21(DE3) high expression strain, induced, lysed, purified with Ni-NTA columns (Error! Reference source not found.), and protein concentrations quantified using calibrated Nanodrop readings (Error! Reference source not found.). Past studies have shown that

glnA aggregates in the presence of Mn and Zn[27,28], however a surprising observation is that both pyrG and glnA aggregate in response to a variety of other divalent metals such as Mn, Co, Ni, Cu, and Zn (**Figure 1c;** Error! Reference source not found.). At 100 μM pyrG or glnA added to 1 mM of Mn, Co, Ni, Cu, Zn, Cd, Hg, or Pb visibly turned cloudy and eventual formed aggregates which could be spun down and pelleted. To test whether 1 mM of metal was causing protein denaturation or non-specific aggregation, a control of 100 μM BSA was added to 1 mM of the same metals, and no visible opaqueness formed, except for Pb which naturally formed hydroxides over time (**Error! Reference source not found.**). Alkaline and alkaline earth metals (Na, Mg, etc.) were not observed to induce aggregation.

The aggregating tendencies of pyrG can be observed visually; pyrG forms spontaneous aggregates at room temperature, and further examination under transmission electron microscopy (TEM) show haphazard connections rather than its canonical tetrameric structure. Under high resolution TEM (HRTEM), pyrG was observed to not regularly form its tetrameric structures, but instead bundle as monomers. In the presence of metals pyrG forms a combination of tetrameric but more often irregular aggregated structures, and at the extreme aggregates into amorphous supramolecular complexes (**Figure 1d**; Error! Reference source not found.**a**). On the other hand, purified glnA examined under HRTEM forms uniform dodecamer structures. Upon addition of metal these structures aggregate into rods, with each face stacking on top of one another. At the extreme these stacked rods aggregate into bundles to form similar amorphous supramolecular complexes like pyrG (**Figure 1e**; Error! Reference source not found.**b**). Given these unique behaviors of metal-induced aggregation and the ability to physically pellet aggregates, this study utilizes the metal responsiveness of pyrG and glnA to chelate heavy metals away from waters for bioremediation applications.

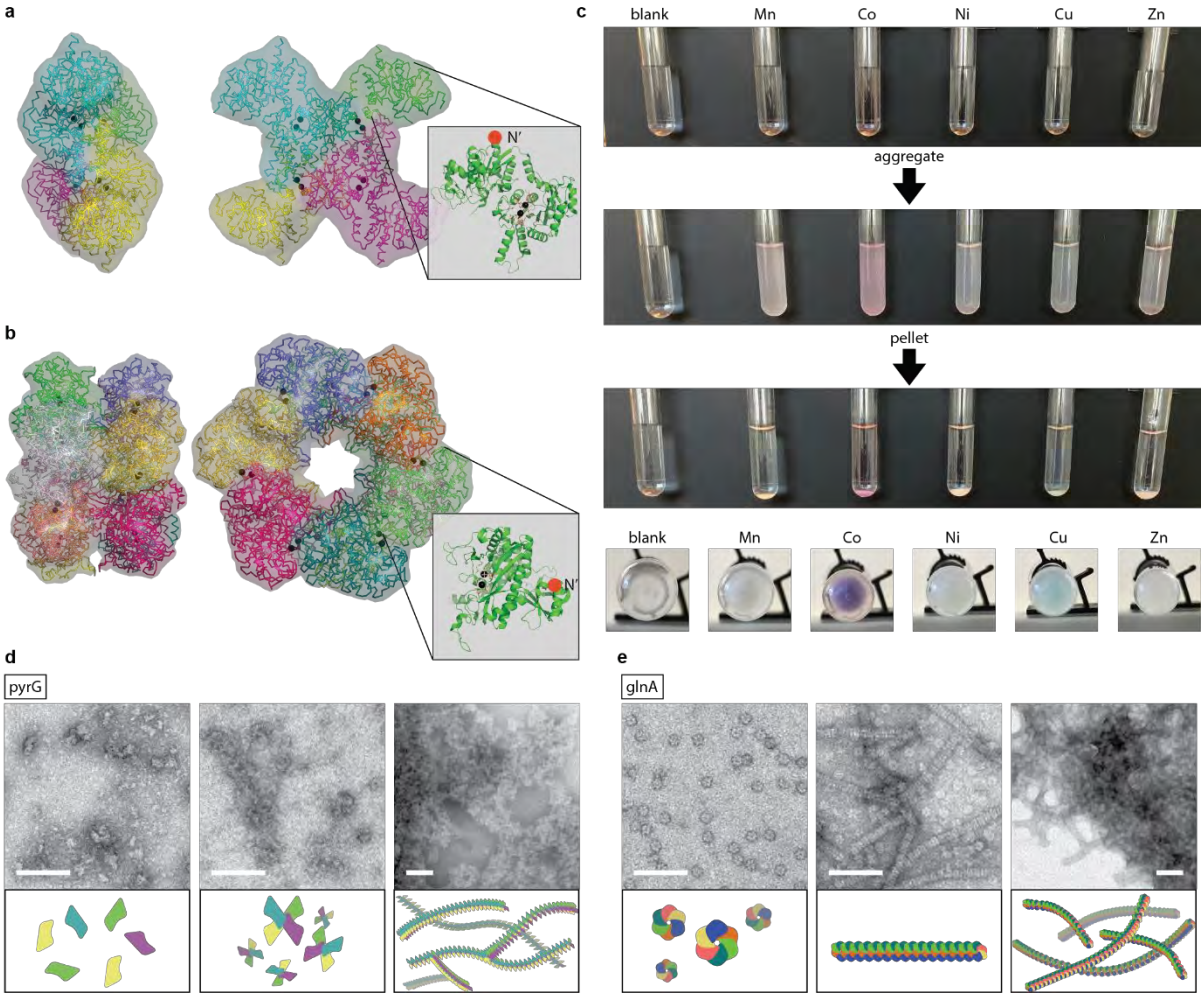


Figure 1 | Using pyrG and glnA as aggregation agents for heavy metal removal. a) pyrG representative crystal structure 1S1M is shown, which contains 4 identical monomers shaped into an 'X'. Each monomer contains 2 divalent metal binding pockets. The N'terminus is highlighted to show the region in which metal binding appendages can be attached to further enhance metal binding. **b**) A similar structure is shown for glnA using representative crystal structure 1FPY shaped as two stacked hexagons. Each monomer has two metal binding pockets, and N'terminus is highlighted. **c**) Visual representation of metal binding and aggregation of glnA. Metal and protein concentration were at 10 mM and 500 μM respectively. **d, e**) HRTEM images of pyrG and glnA, respectively at different levels of aggregation in the presence of Zn. Illustrations below the hypothesized structure and formation of these aggregating chains. Scale bars represent 50 nm.

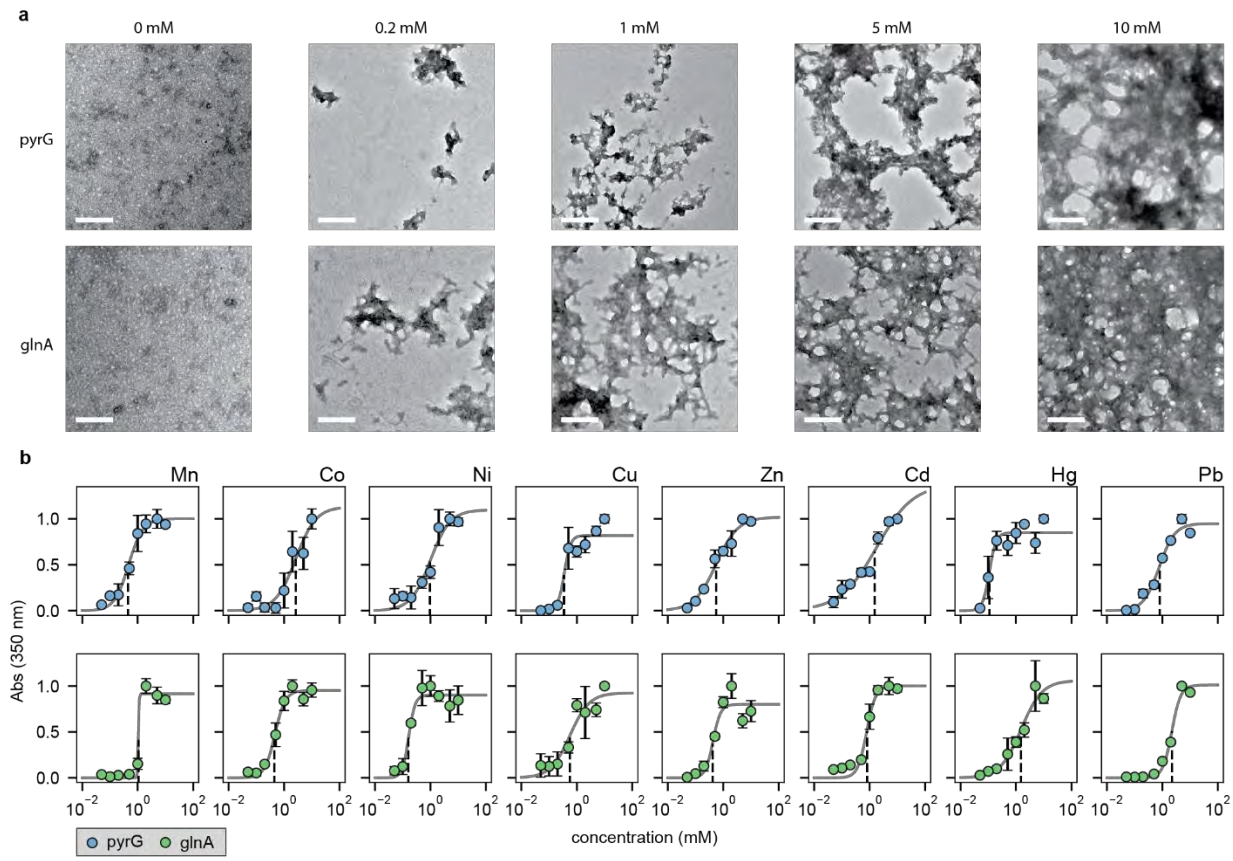
Analyzing metal-induced aggregation of pyrG and glnA

pyrG and glnA metal interactions have been shown to modulate their enzyme activity and are shown in this study to promote protein aggregation[26,27,29]. pyrG and glnA contain 2 metal binding domains per monomer, with binding constants for Mn, Co, and Zn in the low or sub-micromolar range[29,30]. Some researchers have observed that increasing metal concentrations beyond a millimolar retard enzymatic activity, possibly due to aggregation[26,29]. In this study, both pyrG and glnA were observed to form aggregates at 1 mM for Mn, Co, Ni, Cu, Zn, Cd, Hg, and Pb. To quantitatively measure aggregation intensity without having to repeatedly examine structures under TEM, aggregates were measured spectroscopically at 350 nm which was a wavelength that produced the highest signal to noise ratio of aggregates to non-aggregated protein and showed a linear dependence on aggregation density in solution (Error! Reference source not found.). pyrG and glnA titrated with various concentrations of metals show metal-dependent aggregation following typical binding kinetics. Examined under TEM, the density of the protein network increases with increasing metal concentrations (**Figure 2a**). The study was performed again by quantifying aggregates at 350 nm to model a binding (in this case, aggregate) curve. Absorbance readings were normalized and plotted against metal concentrations to form a titration curve fitted with a Hills function. The magnitude of aggregation (*A*), the metal concentration at which aggregation was half-max *A/2* (K_D), and Hill coefficient (*n*) were fitted and tabulated (**Table 1**). The maximum aggregation intensity at 10 mM metal for pyrG followed Pb ~ Cd ~ Cu > Hg > Zn ~ Ni > Co > Mn, whereas for glnA was Pb > Mn ~ Cu ~ Cd > Zn > Co > Ni > Hg. Aggregation sensitivity ranked by K_D (smallest value to highest) for pyrG followed Hg > Cu > Mn ~ Zn > Pb > Cd ~ Ni ~ Co whereas for glnA was Ni > Co ~ Zn > Cu > Cd > Mn > Hg > Pb **(Figure 2b**; **Table 1**). The Hill coefficient for pyrG and glnA for most metals exceeded 1, suggesting positive cooperativity for aggregate formation.

		Mn	Co	Ni	Cu	Zn	Cd	Hg	Pb
pyrG	*A*	0.1 ± 0.01	0.15 ± 0.07	0.20 ± 0.08	0.37 ± 0.04	0.20 ± 0.01	0.395 ± 0.05	0.243 ± 0.01	0.383 ± 0.01
	K_D	0.5 ± 0.01	4.66 ± 0.43	3.7 ± 0.4	0.42 ± 0.2	0.57 ± 0.03	2.44 ± 0.12	0.13 ± 0.05	0.76 ± 0.04
	n	1.58 ± 0.21	12.43 ± 1.6	1.67 ± 0.87	11.06 ± 1.1	1.1 ± 0.06	0.66 ± 0.15	4.31 ± 1.96	1.59 ± 0.22
glnA	*A*	0.22 ± 0.01	0.15 ± 0.01	0.12 ± 0.01	0.21 ± 0.03	0.16 ± 0.002	0.18 ± 0.01	0.08 ± 0.01	1.37 ± 0.1
	K_D	1.05 ± 0.02	0.48 ± 0.06	0.17 ± 0.01	0.61 ± 0.14	0.42 ± 0.05	0.8 ± 0.09	2.11 ± 0.86	2.18 ± 0.04
	n	32.86 ± 3.15	3.32 ± 1.76	4.57 ± 1.81	9.93 ± 1.3	3.25 ± 0.5	2.96 ± 0.76	1.33 ± 0.6	2.7 ± 0.14

Table 1 | Values for maximum aggregation intensity (*A*; measured at 350 nm), aggregation K_D, and fitted cooperativity coefficient (*n*) for pyrG and glnA for the various metals studied. Coefficients were fitted from data collected in Figure 2b.

Metal aggregation was also found to be reversible. Different dilutions of EDTA was added to samples aggregated with 1 mM of metal. Examination under TEM showed that 1-10 mM of EDTA reversed aggregation and released pyrG and glnA into smaller aggregates or individual monomers (Error! Reference source not found.**a**). Examining these effects spectroscopically at 350 nm, a steep reduction in absorbance readings was observed after 1 mM, or at an equal molar ratio of metal to EDTA (Error! Reference source not found.**b**). This finding suggest that pyrG and glnA require some type of metal binding to induce aggregation, and aggregation is reversible upon metal removal or competition.



188

189 **Figure 2 | Measuring metal induced aggregation responsiveness and intensity for pyrG and glnA. a**)
190 Top/bottom row are pyrG and glnA respectively. Columns represent concentrations of Zn which was used to induce
191 aggregation, with progressively higher concentrations leading to higher levels of aggregation. Scale bars represent
192 200 nm for all images. **b**) pyrG and glnA were titrated with metals to measure the level of aggregation as a function
193 of metal concentration. Aggregation was quantitatively measured using 350 absorbance readings (Error! Reference
194 source not found.). For all data, the mean ± s.d. of three replicates are shown.

195

196 **Metal removal via protein-metal aggregation**

197 Aside from understanding the aggregating behaviors of pyrG and glnA, the main objective of
198 this work was to chelate and sediment metals upon induced aggregation. To increase binding
199 capacities per pyrG and glnA complex, their surfaces were decorated with metal binding
200 domains. The exposed N'-terminus was used to fuse several metal binding domains (**Figure**
201 **1a,b**), namely the 6xHis tag, a well-known metal binding motif[31]. Further mention of pyrG or

glnA have an appended N'-terminus 6xHis tag, unless otherwise noted to be cleaved to measure un-modified pyrG and glnA as controls. All experiments used 100 μM of protein and 1 mM metal for aggregation and metal removal studies.

Two distinct factors contribute to the level of metal uptake. The first is the degree of metal aggregation which physically assembles the protein and allow it to sediment out from solution. The second is the degree of metal binding to the protein and metal binding domain, in this case the 6xHis tag. Without the former, no matter how strong the metal binding the protein-metal complex will remain in solution if it does not aggregate and sediment. On the other hand, if the metal binding domain is insensitive to the present metal, no amount of aggregation will chelate the metal away. Therefore, pyrG and glnA have different metal removal profiles when analyzing removal of Mn, Co, Ni, Cu, Zn, Cd, Hg, and Pb (**Figure 3a,b**). For example, pyrG does not aggregate strongly with Mn whereas glnA does, therefore metal removal for Mn is almost 4 times higher for glnA ($p < .05$); however, the opposite is true for Hg where aggregation is more pronounced for pyrG, corresponding to 4 times more metal removed compared to glnA ($p < .05$). When compared to pyrG and glnA without a 6xHis tag, metal removal is significantly reduced for Ni, Cu, Zn, Cd, Hg and Pb ($p < .01$; except for glnA with Hg, $p < .05$), confirming that the fused metal binding domain increases metal binding capacity and removal. The most amount of metal removed for pyrG were with Cu (463 ± 34 μM), Zn (413 ± 13 μM), and Hg (670 ± 33 μM). Whereas for glnA they were Zn (699 ± 25 μM), Cd (373 ± 28 μM), and Pb (433 ± 51 μM) in the presence of 1 mM metal. Therefore, despite the same 6xHis tag, metal removal was heavily dependent on the degree of aggregation per metal.

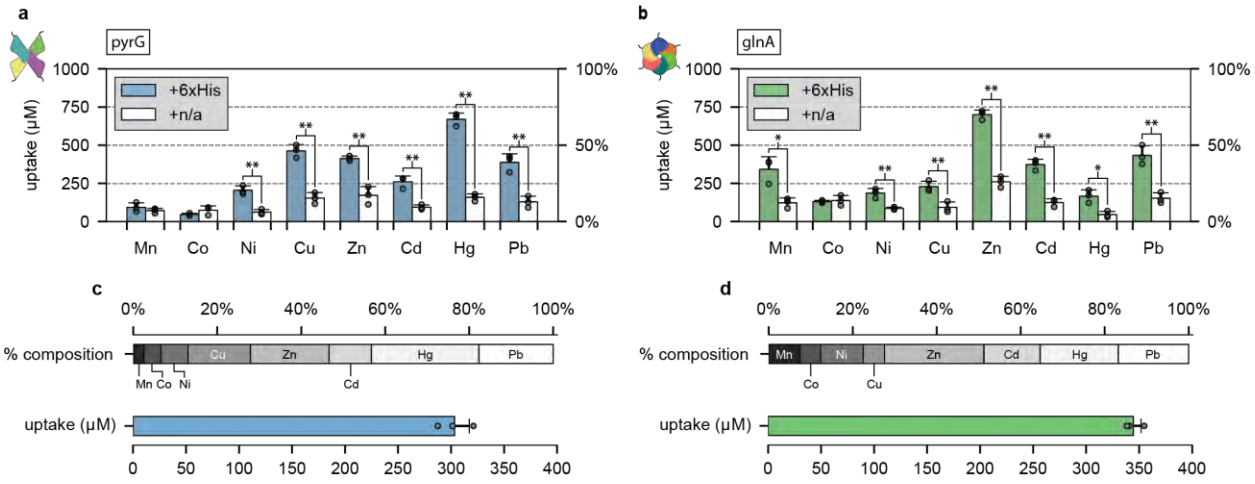


Figure 3 | Metal removal of Mn, Co, Ni, Cu, Zn, Cd, Hg, and Pb were measured for pyrG and glnA. **a, b**) Removal of 1 mM metals were individually measured for pyrG or glnA, respectively. Controls (white bars) are pyrG and glnA cleaved of its 6xHis tag fusion and measured for metal uptake. **c, d**) Metal removal of mixtures of the metals specified (totaling 1 mM, or 125 μM each) were measured for pyrG or glnA, respectively. Top bar represents the percent composition of the metals removed, whereas the bottom bar represents the total amount of metal removed. For all data, the mean $\pm$ s.d. of three replicates are shown.

pyrG and glnA were also tested for mix-metal removal. Mn, Co, Ni, Cu, Zn, Cd, Hg, and Pb were mixed at equal molar to a final metal concentration of 1 mM (125 μM per metal) and mixed with pyrG and glnA. The amount of metal removed for both were approximately >33% (**Figure 3c,d**). For pyrG, the individual metal makeup of what was removed favored Cu, Zn, Cd, Hg, and Pb. Likewise, the same analysis for glnA showed a similar composition of metal removed from the mixed metal source (**Figure 3c,d**; Error! Reference source not found.). Even though pyrG and glnA have different metal aggregating responses, the presence of all 8 metals give pyrG and glnA the opportunity to aggregate freely with the metal that has the largest aggregation effect. Therefore aggregation is no longer an issue and the responsibility for metal removal is now on the 6xHis tag. So the metal removal composition profile appear roughly the same for pyrG and glnA because of the same 6xHis metal-binding domain. For example, to better illustrate this

takeaway glnA has a relatively weak aggregating effect with Hg, despite a relatively strong affinity for 6xHis (**Figure 3b**). However, if the interest is to remove Hg, then glnA can be co-mixed with Zn, a metal that induces a strong aggregating effect in glnA (Error! Reference source not found.). The outcome is an increase in Hg removal by encouraging aggregation with another metal, Zn, allowing the 6xHis tag to now pull down both Hg and Zn. The results shown in **Figure 3c,d** is an extension of this example, where the composition of metal removed is more dependent on the 6xHis tag since protein aggregation will inevitably occur in the metal mixture.

Heavy metal removal is ultimately a numbers game, and to understand this metal removal system pyrG and glnA's stoichiometry's were analyzed to theoretically predict upper and lower bounds on metal removal. pyrG and glnA have two native metal binding sites per monomer (**Figure 1a,b**), and the addition of the 6xHis tag now increases the stoichiometry to at least 3. However, the calculated stoichiometry of pyrG and glnA with their measured uptake give ratios between 3-8 (depending on the metal), higher than expected. Several explanations can be made: the 6xHis tag is binding to multiple metals, the aggregation of proteins may encourage allosteric binding due to the proximity of multiple 6xHis tags leading to positive cooperativity, or the act of aggregation opens other metal binding pockets not observed in the monomer. Data without the 6xHis tag showed removal ratios of almost 1:1, meaning only one of the binding pockets were occupied, a finding that supports past results for pyrG and glnA binding kinetics[29,30]. So the increase in metal removal with the addition of the 6xHis tag suggest that the highly electronegative surface of the protein aggregate disproportionally encourages metal binding than the predicted one 6xHis tag to one metal. Therefore, modifications to the protein surface is an

266 important engineering parameter to design for in order to both enhance metal removal and metal
267 removal specificity.
268
269 **Tuning metal removal by modifying the fused metal-binding domain**
270 Aside from the 6xHis tag, there are many other metal-binding peptides/proteins that can serve as
271 metal-binding accomplices with pyrG or glnA. The remainder of this study focuses on glnA
272 because of its better stability, more defined macrostructure as examined under TEM, and the
273 higher metal-protein binding stoichiometry per macromolecule compared to pyrG (12 compared
274 to 4).
275
276 A family of metal-binding proteins used in a variety of cells for metal detoxification are
277 metallothioneins (MTs). MTs are cysteine-rich, low molecular weight proteins with high metal
278 affinity for common toxic metals such as copper, cadmium, and mercury[8]. Many of the isoforms
279 of MTs have been studied in plants, and because *E. coli* do not natively have MTs, the MT1A
280 (#P43392) from *A. thaliana* was codon-optimized and fused to the N'-terminus of glnA replacing
281 the 6xHis tag. Performing an identical experiment with the 6xHis tag variant, the profile of
282 metals removed skewed more towards Cu, Zn, Cd, Hg. The uptake of Cu, Zn, Cd, and Hg
283 increased by 406 ± 92 μM ($p < .01$), 110 ± 13 μM ($p < .05$), 169 ± 62 μM ($p < .01$), and $147 \pm$
284 53 μM ($p < .05$), respectively, in comparison to using a 6xHis tag (**Figure 4a**). In a multi-metal
285 uptake experiment the enhanced metal removal is more apparent, with total metal removal
286 exceeding 50% and a large composition of the metal removed being Cu, Zn, Cd and Hg (**Figure**
287 **4b**).
288

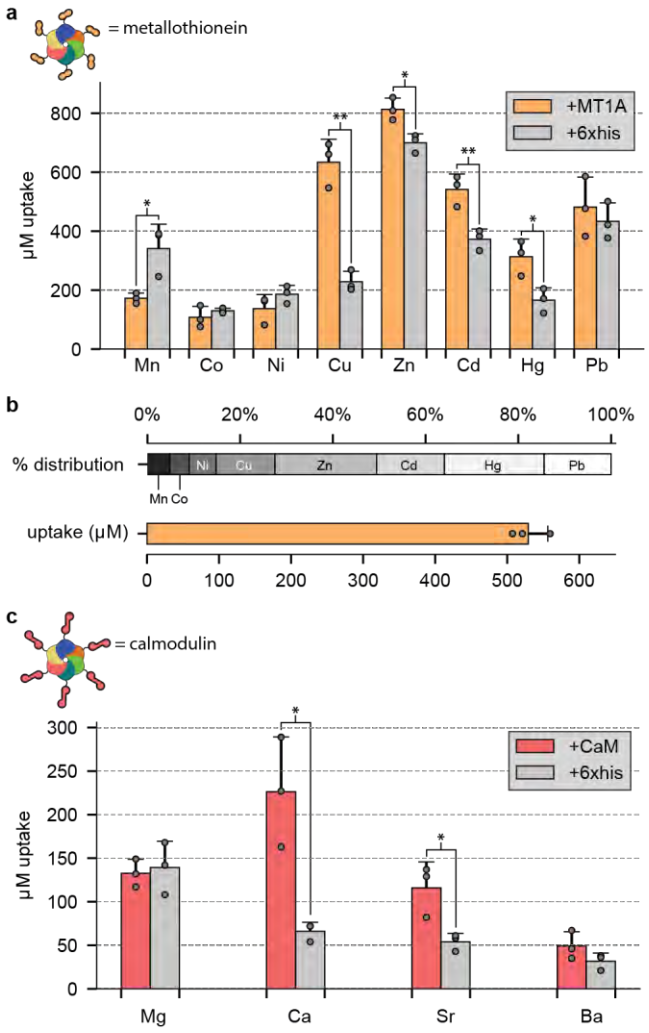


Figure 4 | Substituting the 6xHis tag with plant MT1A or yeast calmodulin alters metal binding preference and metal removal capability. a) The same metal removal experiment was performed for +MT1 as was with glnA with a 6xHis tag. Individual metal preference skews more towards copper, zinc, cadmium, and mercury ($p < .05$). **b**) Likewise, the same metal mixture removal experiment was performed, where the top bar represents the percent composition of removed metals, and the bottom bar represents total metal removed. **c**) +CaM were aggregated with Zn and tested for Mg, Ca, Sr, and Ba removal. For all data, the mean ± s.d. of three replicates are shown.

A similar experiment can be performed for alkaline-earth metals such as calcium by exchanging the 6xHis/MT1 fusion for an alkaline-earth sensitive binder. Although elements such as magnesium and calcium are not dangerously harmful when compared to cadmium or mercury, they are still harmful for remediation processes as these metals often precipitate which erode

infrastructure, calcify around sensitive plumbing, and add more background metals that can effect remediation processes like ion-exchange[32]. A commonly used protein for calcium binding, calmodulin (CaM) was used to test whether the metal removal profile can be shifted to alkaline-earth metals. Yeast CaM (#P06787) was isolated from the genome of *S. Cerevisiae* W303α strains and cloned into glnA to replace the N'-terminus 6xHis sequence. Unfortunately, the alkaline-earth metals do not induce aggregation, therefore 1 mM Zn was co-mixed with either 1 mM Mg, Ca, Sr, or Ba to promote sedimentation and metal removal. Replacing 6xHis for CaM showed increased calcium removal of 226 ± 51 μM compared to 66 ± 8 μM of glnA with 6xHis ($p < .01$), and a slight increase in Sr removal of 116 ± 30 μM compared to 54 ± 7.8 μM of glnA with 6xHis ($p < .05$) (**Figure 4c**). The metal removal profile loosely corresponds to CaM native preference for Ca, followed by Mg and Sr[33].

Improving sedimentation metal-protein complexes and metal recovery

The metal-induced precipitation of glnA offers a natural self-filtering mechanism via protein aggregation and sedimentation. However, rather than waiting several hours, or forcibly pelleting aggregates using a centrifuge, sedimentation of glnA can be enhanced by binding aggregated complexes to denser anchors such as beads, or even cells. The effect on sedimentation was tested by binding aggregated glnA with agarose or magnetic beads. In addition, a fully biological route was to display glnA monomers via yeast display and allowing aggregates to grow and anchor onto the yeast cell surface. For bead attachment, glnA was modified to include expression of a 3xFlag tag at the N'-terminus in conjunction with the 6xHis tag. Agarose or magnetic beads were functionalized with anti-flag antibodies to recognize the 3xFlag tag and pull down the aggregates. For the yeast display condition, EBY100 strains were transformed with the pYD1

yeast display vector expressing a glnA monomer (no appended His tag) (Error! Reference source not found.). glnA aggregation was performed together with yeast displaying cells so aggregates could nucleate onto the yeast cell surface. In this setup the 3xFlag tag was not added to the glnA protein.

Measurements of sedimentation was determined by measuring protein concentration at the liquid's mid-height. When placed in cuvettes at 4 mL (mid-height being at 2 mL), 100 μM glnA aggregated with 1 mM of Cd took several hours to sediment. At roughly 5 hours almost half of the aggregated content was beneath 50% of the sampled liquid height (**Figure 5a**). When aggregates were mixed with agarose or magnetic beads, sedimentation was much quicker. Using agarose beads, it took 1 minute for 50% of the aggregates to sediment beneath the sampled liquid height. Whereas, with magnetic beads and using an external magnet for pull down, sedimentation was much more dramatic only requiring < 5 seconds for the same amount of protein removal (**Figure 5a,b**). Alternatively, glnA and yeast displaying glnA monomers were co-precipitated for 1 hour with 1 mM Cd to allow aggregates to bind onto the yeast surface. Because yeast are only slightly denser than water (> 1.1 g/mL), it took more than an 1 hour for protein-bound yeast complexes to reach the same level of sedimentation when compared to agarose or magnetic beads (**Figure 5a,b**). However, when glnA was not modified with a 3XFlag tag, or yeast displaying an empty vector were mixed, sedimentation was largely unaffected and still required several hours for aggregate removal (Error! Reference source not found.).

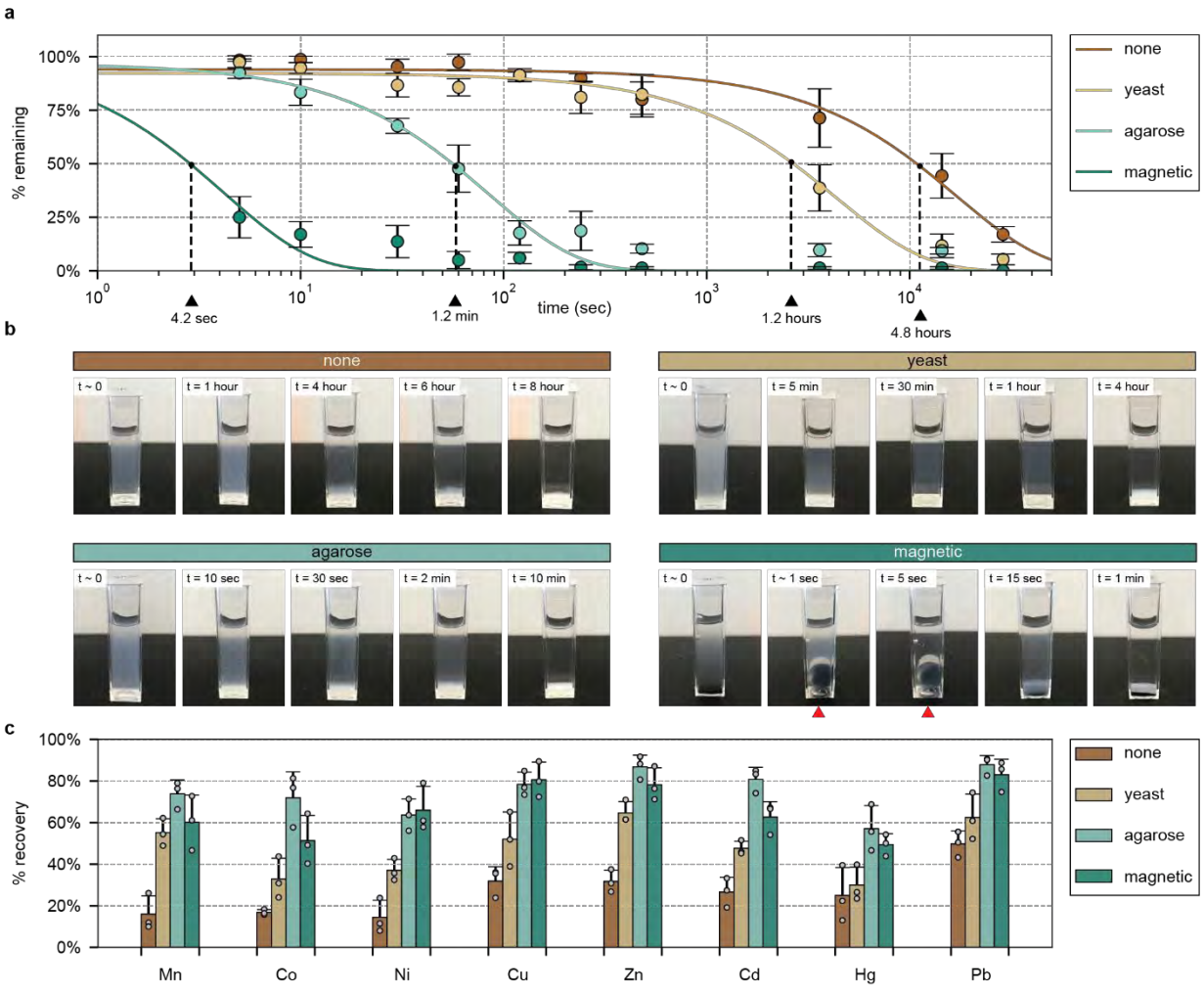


Figure 5 | glnA fused with a 3XFlag tag, or using yeast displaying glnA monomers, improve sedimentation rates and metal recovery. a) glnA+3XFlag was aggregated with 1 mM Cd and later mixed with anti-flag agarose or anti-flag magnetic beads, or glnA was mixed with yeast displaying glnA monomers. Magnetic beads had the quickest pull down of aggregates, followed by agarose beads, then yeast. **b**) Visual representation of aggregation sedimentation with yeast display, agarose, or magnetic beads. Red arrows for the magnetic beads indicate moments when an external magnet was used. **c**) Mixtures were allowed to sediment after 1 hour in which the sedimented pellet was isolated, washed, and extracted of metals using EDTA. Bars represent the percent metals recovered relative to the amount of metals removed by the protein aggregates. For all data, the mean ± s.d. of three replicates are shown.

Metal recovery was also tested using the various sedimentation strategies. Before sedimentation, the amount of metal bound to glnA was measured. For the yeast condition, glnA and yeast displaying glnA were mixed together during aggregation, complexes were gently pelleted, and

359 the supernatant was sampled to calculate the amount of metal captured in the complex. For both
360 cases, after measurement the mixtures were then resuspended to carry out the sedimentation and
361 metal recovery experiment. After resuspension, the mixture was allowed to sit for 10 minutes at
362 which the supernatant was fully removed and whatever was sedimented was recovered. The
363 sediment was transferred and washed 2X in ddH_2O before resuspending in 10 mM EDTA. The
364 amount of freed metals were measured and compared against the initial measurement of metals
365 bound to the aggregate before the sedimentation experiment. This value gave metal percent
366 recovery given the 10 minutes allowed for sedimentation. Factors that affect recovery were the
367 rate of sedimentation within the 10-minute timeframe, and the binding strength between the
368 aggregates and anchors during the wash steps. Without any anchor, the recovery for just glnA
369 was below 10% for all metals (**Figure 5c**). For yeast it was observed that flakes of aggregates
370 would dissociate during the wash steps; this was not seen for the agarose or magnetic bead
371 samples. Aggregates with yeast displayed monomers had recoveries between 15-40%. When
372 using beads, the agarose and magnetic anchors were quicker to sediment thereby recovering
373 more aggregates which corresponded to a higher percent recovery of Cu, Zn, Cd and Pb which
374 were between 60-80%. The high recoveries could also be associated to the tight binding of the
375 3XFlag tag between glnA and the functionalized antibodies on the beads. A takeaway from these
376 experiments is that not only can glnA be modified to increase metal removal and metal removal
377 preferences using appended metal binding domains, but further modifications can lead to other
378 functionality such as attachment to beads or cells for improved sedimentation and recovery.
379

Discussion

Biological enzymes previously discovered to have unique structural and enzymatic roles in cells can be re-purposed for bioremediation and water cleaning strategies. The appeal of using such proteins is their controllable aggregation behavior and the subsequent solid sediment that forms allowing for physical handling and removal from contaminated waters. This study shows that pyrG and glnA are responsive to a variety of divalent metals such as Mn, Co, Ni, Cu, Zn, Cd, Hg, and Pb and aggregate within the submillimolar to millimolar range. In addition, pyrG and glnA surfaces are easily modified and appended with metal chelating agents such that metal removal is physically captured in the aggregated sediment. In this perspective, pyrG and glnA act as vehicles in which their surfaces can be genetically modified to introduce metal chelating agents for heavy metal clean up. For example, adding a 6xHis tag to pyrG and glnA has broad metal specificity for most metals tested, and increased metal removal for Cu, Zn, Cd, Hg and Pb compared to pyrG and glnA without. The 6xHis tag can be substituted for other small peptide/protein metal binders such as a metallothionein or calmodulin which alter metal removal profiles. For example, the metallothionein MT1A from *A. thaliana* increased overall metal removal for Cu, Zn, Cd, and Hg compared to the 6xHis tag. More so, metal preferences were entirely altered from transition metals to alkaline-earth metals using calmodulin from yeast. Appending calmodulin showed preferential calcium removal, whereas the 6xHistag was insensitive. The ability to engineer metal removal preferences from heavy metals to alkaline metals provides another avenue for water cleaning which is for water softening, a frequent problem in municipal drinking waters and plumbing infrastructure.

Protein anchor glnA can be further modified to contain binding tags that can adhere to denser anchors such as agarose or magnetic beads for faster sedimentation. Likewise, a fully biological method used yeast display to anchor aggregates onto the yeast surface. With either beads or yeast the time for sedimentation reduced from hours (glnA alone) to minutes (agarose) or even seconds (magnetic beads). Once sedimented, the pull-downed aggregate-metal complex were dissociated to release the bound metal for recovery purposes. Overall, improved sedimentation due to agarose and magnetic beads correlated to improve metal recovery.

Different metal binding domains, repeat of domains, or combinations of several binders on a single protein can potentially improve protein-based water removal capacities and tailor metal removal preferences. For example, repeats of the 6xHis tag can be added to each glnA monomer to potentially increase the number of metal binders, and hence improve removal capacities. Alternatively, several proteins with different metal binding preferences can be fused on one chain to customize metal removal profiles. For example, fusing MT1a with CaM can potentially remove both transition and alkaline-earth metals. In addition, a variety of tags can be added for additional functionality; in this work, they were denser beads to improve sedimentation rates. However, future possibilities could be to anchor pyrG or glnA on other platforms using well-established protein binding strategies (Flag, Streptavidin, SpyTag, etc.). Several interesting platforms would be biofilms, hydrogels, or other semi-porous matrixes that can harbor metal removing proteins like pyrG or glnA. More possibilities not mentioned (or thought of) here can be easily designed by appropriately modifying the surfaces of pyrG or glnA and combining it with the platform of choice. What may limit the length or complexity of appendages would be a disruption to pyrG or glnA protein expression and folding. However, this concern of over

425 burdening the protein is true for most protein-fusion expression systems and requires fine tuning
426 of expression conditions and purification protocol.
427
428 The identity of the protein anchor is not as important as its ability to aggregate and sediment out
429 of solution. There has been much work in the biological community to identify the growing
430 number of enzymes that behave like pyrG and glnA, and so far a list of 33 proteins discovered in
431 yeast may have similar properties[24], and many more may exist in other organisms[22,23]. Although
432 not exhaustively tested here, there is a possibility that each protein behaves differently in either
433 responsiveness or sensitivity towards aggregation, and these behaviors can be uniquely
434 capitalized for different heavy metal removal conditions. More so, protein stability, stability with
435 respects to surface modifications, and controlled aggregation are desired behaviors of the host
436 protein. Future steps would be to understand mechanistically and biochemically how aggregation
437 can be induced and controlled in this set of discovered proteins. If these underlying mechanisms
438 can be uncovered, more engineering can be done to improve the aggregation sensitivity and
439 density of aggregates for improved metal removal, sedimentation, and recovery.
440
441 The goal of this work was to biologically mimic ion-exchange by recapitulating its underlying
442 mechanism in proteins. The first was to use aggregating proteins to physically mimic the resin
443 bed. The second was to fuse metal binding peptides/proteins onto these aggregates for targeted
444 metal removal, a process similar to functionalizing resins with strong and weak metal
445 exchangers. However, unlike ion-exchange these proteins are autonomously produced in cells
446 and do not require complex chemical synthesis normally used when creating resins. In addition,
447 protein modifications are becoming increasingly easier to perform on the genetic level given the

maturing technologies in genetic and protein engineering. This ever-growing synthetic biology toolkit opens a range of customizability when designing aggregating proteins and metal binders. Finally, using a fully biological method to create resin like ion-exchangers is appealing because protein production can be more economic and environmentally sustainable than manufacturing its synthetic ion-exchange counterpart. Proteins focused here are naturally available, require little chemical processing, and are robustly produced in *E. coli*. Much like the antibody market has grown to be a mass producer of proteins[34], the same infrastructure and technology can be leveraged to produce proteins for applications in clean water technology. As biological engineers, the hope is to take advantage of current biological technologies to efficiently and sustainably solve the waste water crisis, and more so, convince others that many more solutions may exist with further exploration at the intersection of biology and technology.

Methods

Gene isolation and plasmid construction

Genomic DNA from *E. coli* DH5α (NEB) was isolated using DNA Fungal/Bacterial Miniprep Kit (Zymo). Gene sequences of pyrG and glnA were retrieved via Uniprot and used to create primers (IDT) for Gibson assembly into pET28c(+) IPTG-inducible vector (Error! Reference source not found.). 5' primer contained an overhang adding a TEV protease site (ENLYFQS) downstream of the pET28c(+) 6xHis tag before the gene insertion. Primers were used to amplify the appropriate genes using PCR with the Q5 polymerase (NEB). All PCR products mentioned were examined under a gel imager (AlphaImager 2200) and cleaned (Promega) before performing subsequent cloning steps. Products were then assembled into linearized pET28c(+) using the HiFi 2X assembly master mix (NEB). Assembled constructs were transformed into NEBα cells (NEB), plated onto 1X kanamycin (50 mg/L) LB plates, picked and miniprepped (Promega), and sequenced to confirm proper gene insertions (Quintara Bio). Confirmed sequences were then transformed into BL21(DE3) (Agilent) for protein expression.

The pET28c(+)-glnA vector was further modified by fusing calmodulin (CaM) or *A. thaliana* metallothionein 1a (MT1a) before glnA. The DNA sequence for CaM was isolated from the yeast genome (also purified using the DNA Fungal/Bacterial Miniprep Kit), whereas the MT1a sequence was codon-optimized and synthesized (GenScript). Primers were created for Gibson assembly into the pET28c(+)-glnA vector (Error! Reference source not found.) following the protocol already described above.

Alternatively, pET28c(+)-glnA was modified by appending a 3XFlag tag before glnA. Primers were designed to anneal before the glnA sequence containing overhangs carrying the 3XFlag sequence (DYKDHDGDYKDHDIDYKDDDDK) (Error! Reference source not found.). The linear PCR product was re-circularized by adding T4 polynucleotide kinase with T4 ligase in 1X T4 buffer (NEB) before bacterial transformation.

The pyrG and glnA were also used for yeast display in the pYD1 vector. Primers were designed with a 5' NheI and 3' BamHI restriction cut sites (Error! Reference source not found.) and isolated using the protocol already described above. The pYD1 vector was linearlized by cutting with NheI-HF and BamHI-HF restriction enzyme (NEB). pyrG or glnA were ligated into linearized pYD1 using T4 ligase and then used for subsequent transformations.

Protein purification

BL21(DE3) cells were grown in 10 mL starter cultures overnight. The following day cultures were diluted by 1:100 in small scale (100 mL) or large scale (1 L) cultures. Cultures were grown for 5-6 hours at 30ºC before induction with 1X IPTG (1 mM; GoldBio). Cultures were induced for 16 hours before harvesting by pelleting cultures at 5000xg for 15 min. Pelleted cultures were stored in -80ºC for 1 day before preparing protein samples.

For 100 mL of pelleted culture, 10 mL of 1X Bugbuster (Millipore Sigma) was used for cell lysis. For each mL of Bugbuster, 1 μL of lysonase bioprocessing reagent (Millipore Sigma), 1 μL of DNase (NEB), and 10 μL of 100X Halt Protease Inhibitor Cocktail (Thermo) were added. Resuspended cultures were gently agitated on a vortexer for 15-30 minutes until solutions became

clear. Lysed solutions were spun down at 16,000xg for 20 minutes at 4ºC to remove debris and inclusion bodies. During this process Ni-NTA resin (Qiagen) was prepared by taking 2 mL for every 100 mL of pelleted culture and equilibrating with 25 mL 1X Ni-NTA Buffer (Millipore Sigma). Resins were spun down at 1000xg for 5 min and supernatant removed. Supernatant was removed and filtered through a 0.45 μm syringe filter directly onto equilibrated Ni-NTA resin. Mixtures were gently rocked for at least 2 hours at 4ºC.

Mixtures were poured into disposable Econo-Pac chromatography columns (10 or 20 mL, depending on the scale; Biorad). Columns were washed 3 times with 5X column volumes (1 mL NTA resin equals 5 mL of wash buffer) of 1X Ni-NTA wash buffer (Millipore Sigma). Protein was then eluted with 2X column volume of 1X Ni-NTA elution buffer (Millipore Sigma) separated into 4 fractions (1 mL of NTA resin equals 0.5 mL of elution buffer, repeated 4 times).

Eluted proteins were desalted using Sephadex G-25 PD-10 desalting columns (GE). Columns were first equilibrated with 1X TBS (100 mM Tris in 150 mM NaCl; Rockland Inc.). 2.5 mL of eluted protein were flowed through, followed by 3.5 mL of 1X TBS which was collected. For protein samples that needed to be cleaved of the 6xHis tag, namely pyrG and glnA controls, and glnA with CaM and MT1A add-ons, samples were instead dialyzed. TEV protease (Sigma) was added to eluted samples at 1:100 (w/w), typically equating to 1:500 (v/v). Samples were fitted in 10K MWCO Snakeskin dialysis membrane (Thermo) and dialyzed against 5 L of 1X TBS for 2 days at 4ºC. Dialysis buffer was exchanged every 12 hours. Dialyzed samples were then added back to Ni-NTA columns to remove cleaved 6xHis tag, uncleaved protein, and TEV itself.

Proteins that were below 1 mg/mL were concentrated using Amicon centrifugal filters (Millipore Sigma) with 10K cutoffs. Filters were spun at 4000xg at increments of 5 minutes in a swinging-bucket rotor to reduce sample volume and increase protein concentration.

Protein expression was examined using SDS-PAGE and coomassie staining. Protein samples were diluted to 0.5 mg in 20 μL sample volume containing 1X LDS buffer and 1X denaturation buffer (Thermo). Samples were boiled at 70ºC for 10 min and loaded onto pre-cast Bolt 4-12% Bis Tris gels (Thermo). Gels were ran at 200 V for 30 min in 1X MES buffer (Thermo) and visualized using PageBlue protein stain (Thermo). Band sizes were compared against the PageRuler pre-stained protein ladder (Thermo). pyrG, glnA, glnA+CaM, glnA+MT1, and glnA+3XFlag gave single and clear bands. +CaM showed a strong band at the appropriate size, and another fainter band where glnA is normally found, suggesting that the fusion protein was cleaved during the expression or purification process. However, the alkaline metal removal studies for glnA+CaM should not be effected by the presence of background glnA, as glnA has little sensitivity for the metals tested (Mg, Ca, Sr and Ba).

Protein quantification

Protein concentrations were routinely checked on the Nanodrop. However, to better calculate protein concentrations a relationship between the higher resolution Pierce 660 nm protein assay (Thermo) and Nanodrop readings was performed. Serial dilutions of pyrG, glnA, glnA+CaM, glnA+MT1A, glnA+3XFlag starting at 2 mg/mL were read on both Nanodrop and 660 nm assay. A line of best fit ($y = m \cdot x + b$) was constructed to give a relationship between the 660 nm reading and Nanodrop, i.e. $[660] = \mathrm{m_{nano}}[\mathrm{Nano}] + \mathrm{b_{nano}}$. The Pierce protein assay was then

used to create a calibration curve against BSA standards (Thermo), which gave a calibration curve of protein concentration (mg/mL) versus 660 nm, i.e. $[\mathrm{mg/mL}] = \mathrm{m}_{660}[660] + \mathrm{b}_{660}$. Substituting the relationship between Nanodrop and the Pierce 660 nm protein readings, the final equation relating Nanodrop to the BSA calibrated protein concentration reading of mg/mL was $[\mathrm{mg/mL}] = \mathrm{m}_{660} \cdot \mathrm{m}_{\mathrm{nano}}[\mathrm{Nano}] + \mathrm{m}_{660} \cdot \mathrm{b}_{\mathrm{nano}}$. This new line of best fit was used to correlate readings from the Nanodrop to mg/mL as determined by the more reliable Pierce 660 nm protein assay (Error! Reference source not found.).

Protein concentrations were calculated using the monomer molecular weight. So, 62.5 kDa for pyrG and 54 kDa for glnA.

Aggregation quantification using absorbance measurement

pyrG and glnA without a 6xHis tag were used in these experiments. Purified proteins were aggregated using 10 mM of metals and serially diluted in 96-well plates. Absorbance scans at increments of 5 nm were measured for each sample on a plate reader (Tecan M200 Pro). One wavelength at each dilution was measured and fitted with absorbance (y-axis) versus dilution (x-axis). The wavelength 350 nm gave the best fit between absorbance intensity and protein aggregation, compared to non-aggregated protein which gave negligible signal (**Error! Reference source not found.**).

Proteins at 100 μM were aggregated in 100 μL in a 96-well plate at varying metal concentrations starting at 10 mM. Aggregation was allowed to occur for 1 hour before measuring at 350 nm. Absorbance measurements at 350 nm (y-axis) were plotted against metal concentrations (x-axis)

to fit a Hill equation in order to parametrize the K_D for metal induced aggregation, and intensity of aggregation (A) (i.e. intensity) and Hill coefficient (n) (**Table 1**).

Testing reversibility of protein aggregation was performed by aggregating proteins at 1 mM metal for 1 hour. Resuspended aggregates were aliquoted in 100 μL in a 96-well plate at a final concentration of 100 μM. In each well a different concentration of EDTA was added, with the highest at 10 mM and serially diluted by factors of 10. EDTA and protein aggregate were mixed for 10 minutes. The plate was then measured at 350 nm to measure intensity of protein aggregation after metal removal due to EDTA.

Metal uptake experiments

Protein aggregates were first formed by mixing 100 μM of protein with 1 mM of metal. The reaction was allowed to occur for 1 hour before spinning down samples at 10,000xg for 5 min at 4^{o}C. The supernatant was collected and diluted 1:10 in 3% HNO_3 solution for ICP measurement. An ICP-OES (Agilent 5100) was used to measure metal concentrations in the supernatant. Standards were made from ICP-quality metal stocks (Fluka). In addition, samples with metal added but no protein were measured to test for natural metal precipitation or non-specific metal binding onto the sample tubes. The only metal with appreciable precipitation was Pb, which naturally forms hydroxides after several minutes. This value was subtracted from the ICP measured in the supernatant in order to adjust for the actual metal removed due to aggregate formation. Finally, the amount of protein captured by the pelleted aggregates was calculated by subtracting the original metal concentration with the adjusted ICP measurement from the supernatant.

For multi-element uptake experiments, the eight studied metals, Mn, Co, Ni, Cu, ZN, Cd, Hg, and Pb were mixed in equimolar ratios with a combined metal concentration of 1 mM (concentration of each metal being 125 μM). The same metal removal experiment was performed as described above. Wavelengths for ICP analysis with the minimal amount of cross-over were Mn (257.610 nm), Co (230.786 nm), Ni (216.555 nm), Cu (327.395 nm), Zn (213.857 nm), Cd (226.502 nm), Hg (194.164 nm), and Pb (220.353 nm). The amount of metals removed was calculated as described above for each metal; the total amount of metal removed being the sum of all calculated values.

For metal uptake experiments of control pyrG and glnA without a 6xHis tag was used. 6xHis tag of purified pyrG and glnA were cleaved using TEV protease and dialysis, and re-purified using Ni-NTA to isolate non-tagged 6xHis proteins. Non-tagged pyrG and glnA controls were used in parallel with tagged proteins during the metal removal studies.

Similarly, glnA+CaM and glnA+MT1 were first cleaved of its 6xHis tag used for protein purification. For glnA+CaM alkaline-earth metal removal experiments, samples were co-mixed with 100 μL of either Mg, Ca, Sr or Ba and 100 μM Zn. Metal removal experiments with glnA+CaM and glnA+MT1 followed the same experimental outline as described above. ICP wavelengths for alkaline-earth measurements were Mg (279.553 nm), Ca (396.847 nm), Sr (407.771 nm), and Ba (233.527 nm).

Transmission Electron Microscopy sample preparation and imaging

All TEM samples were prepared with pyrG and glnA without a 6xHis tag to observe native aggregation effects. 7 μL of sample was removed for transmission electron microscopy (TEM). Samples were dropped on a 400-mesh copper grid coated on carbon film (EMS). Grids were left for 60 seconds before removing excess solution by touching the grid on a kimwipe. 10 μL of negative staining solution phosphotungstic acid at 1% was dropped on the grid and immediately removed with a kimwipe. Another 10 μL of negative stain was immediately dropped after and left for 40 seconds before removing excess liquid. Grids were left to air dry at room temperature for more than 30 minutes.

A FEI Tecnai was used at 120 kV to image protein aggregates with roughly 100 nm resolution. High resolution TEM (HRTEM) was performed on a JOEL 2100 FEG microscope at 200 kV with assistance from Koch's Nanotechnology Core stewards.

Yeast display of CTP and GS monomers

Yeast EBY100 strain were made competent using Frozen-EZ Yeast Transformation II Kit (Zymo) and either stored at -80ºC, or used immediately for transformation. Competent EBY100 was transformed with pYD1 vectors containing pyrG or glnA without a 6xHis tag. Transformed cells were plated on SDCAA media (Teknova) and grown for 1.5-2 days. Transformants were confirmed by extracting DNA via bead-beating with 420-600 μm glass beads (Sigma) and phenol:chloroform (Sigma) extraction and ethanol precipitation. DNA was then amplified with primers flanking the AGA2 and T7 promoter region using PCR, and checked via gel electrophoresis for correct insertion.

Transformed strains were grown overnight in SDCAA media in 30ºC, and diluted 1 to 10 the next day. Cultures continued to grow for another 4-6 hours to mid-log phase before spinning down cultures and resuspending in SGCAA (Teknova). Cultures were induced overnight at room temperature before harvesting. To check for positive expression, induced strains were washed in PBS + 1% BSA and tagged with primary antibodies against the N'-HA tag and the C'-flag tag for 1 hour at 4ºC. Secondary antibodies conjugated with 488 or 647 dye were used against the primary antibodies and stained for 1 hour at 4ºC. Fluorescently tagged cells were then analyzed using flow cytometry on a LSR II (BD Bioscience) and plots of FITC versus PE-Cy5 were analyzed to measure populations of expressing cells with respects to the WT non-expressing samples.

Sedimentation analysis

For agarose and magnetic bead sedimentation studies, a 3xFlag tag was added after the 6xHis tag (e.g. N'-6xHis-3XFlag-glnA-C'). glnA at 100 μM was aggregated with 1 mM of Cd in 5 mL for 1 hour. Samples were then mixed with 100 μL of pure resin pre-equilibrated in ddH_2O of either Anti-Flag M2 Affinity Gel or Anti-Flag M2 Magnetic Beads (Sigma). Aggregated samples with beads were incubated for 1 hour before transferring 4 mL to a fluorimeter cuvette with 4 flush clear sides (Sigma). Samples were left to settle, whereas for the magnetic beads a magnet was dragged to the bottom of the cuvette and left for approximately 5 seconds before removing to magnetically pull down beads with bound proteins.

For yeast display conditions, 1 OD_{600} of expressing EBY100 was mixed with 100 μM glnA and 1 mM of Cd in 5 mL during aggregation. Aggregation was allowed to continue for 1 hour before transferring to a 4 mL fluorimeter cuvette.

For the sedimentation study, samples were collected at the mid-height (e.g. 2 mL for a 4 mL cuvette). 20 μL aliquots were taken at specific time points and quickly washed with 0.2 M glycine HCl pH 3.5 and shaken for 2-5 minutes to dissociate any bound proteins. Tubes were quickly centrifuged and the top 10 μL was measured for protein content using the Pierce 660 nm protein assay. The percent ratio of protein concentration measured per time point versus original concentration (100 μM) was plotted to analyze the sedimentation rate of protein aggregates when bound to agarose, magnetic beads, or yeast displaying glnA. Experimental controls removed the connector between glnA and the denser anchor. For agarose and the magnetic beads glnA without 3xFlag was used. For yeast pull-down an EBY100 strain displaying an empty pYD1 vector was used. The same sedimentation experiment was performed for these controls.

Metal recovery

The same glnA constructs in the sedimentation experiments were used here; glnA with an added 3xFlag tag for agarose and magnetic pull-down, and EBY100 displaying a glnA monomer for yeast pull-down. For the agarose and magnetic samples, 100 μM glnA+3xFlag and 1 mM metal were allowed to aggregate for 1 hour before spinning them down. The supernatant was sampled to calculate the amount of metal captured in the protein-metal complex. This value was set as the initial amount of metal captured. As a control, samples with no protein were sampled to measure the natural precipitation of metal, namely Pb. This value was subtracted from the measurement

initial metal captured measurement to isolate the actual amount of metal removed from the protein-metal complex. Afterwards, aggregates were resuspended and then mixed with 100 μL of pure agarose or magnetic resin pre-equilibrated in ddH_2O. Mixtures were gently shaken for 1 hour before aliquoting 1 mL into Eppendorf tubes. Samples were allowed to settle for 10 minutes, the same time window as in the sedimentation experiments. For samples with magnetic beads, a magnet was dragged to the bottom of the cuvette and left for approximately 5 seconds. Afterwards, the top 900 μL of sample was removed. The remaining volume was diluted with 1 mL of ddH_2O, spun down, and washed once more to remove any metals not bound to protein. 1 mL of 10 mM EDTA was then added to the pellet and mixed for 10 minutes. The tube was spun once more to remove any beads/yeast or protein debri. The supernatant was sampled and measured for metal using ICP. This value was set as the amount of metal recovered. The amount of metal recovered was divided by the initial amount of metal captured to give a recovery percent.

For yeast anchored samples, 100 μM glnA was mixed with 1 mM metal and 1 OD_{600} of induced EBY100 displaying glnA monomers. The mixture was allowed to aggregate for 1 hour before spinning it down. The supernatant was sampled to calculate the amount of metal captured in the protein-yeast-metal complex. This value was set as the initial amount of metal captured. Controls with just induced EBY100 displaying glnA monomers and 100 μM metal were used to measure non-specific binding onto the yeast surface and natural precipitation of metal. This value was subtracted from the initial amount of metal captured to isolate the actual amount of metal removed from just the protein-metal complex. 1 mL of yeast-protein-metal mixture was aliquoted into Eppendorf tubes. The same protocol mentioned with the agarose and magnetic

bead was performed to calculate the amount of metal recoverable. Percent recovery values for all conditions: agarose, magnetic, and yeast along with all metals used in this work were plotted and compared with one another.

Mathematical analysis and plotting

Raw data was collected and stored as csv or excel file formats. Data was imported and analyzed with python using modules such as numpy, pandas, and scipy. Plots were graphed using matplotlib.

Statistical analysis

Statistical parameters including the definitions and values of *n*, SDs, and/or SEs are reported in the figures and corresponding figure legends. When reporting significance, a two-tailed unpaired t-test was performed between observations. Significance was declared when $p < .05$ for all experiments, or as specified in the text. In the figures, * = $p < .05$ and ** = $p < .01$.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding authors on request.

References

1. Fu, F. & Wang, Q. Removal of heavy metal ions from wastewaters: a review. *J. Environ. Manage.* **92**, 407–418 (2011).
2. Üstün, G. E., Solmaz, S. K. A. & Birgül, A. Regeneration of industrial district wastewater using a combination of Fenton process and ion exchange—A case study. *Resour. Conserv. Recycl.* **52**, 425–440 (2007).
3. Motsi, T., Rowson, N. & Simmons, M. Adsorption of heavy metals from acid mine drainage by natural zeolite. *Int. J. Miner. Process.* **92**, 42–48 (2009).
4. Kumar Gupta, V., Ali, I., A. Saleh, T., Nayak, A. & Agarwal, S. Chemical treatment technologies for waste-water recycling—an overview. *RSC Adv.* **2**, 6380–6388 (2012).
5. Amini, A., Kim, Y., Zhang, J., Boyer, T. & Zhang, Q. Environmental and economic sustainability of ion exchange drinking water treatment for organics removal. *J. Clean. Prod.* **104**, 413–421 (2015).
6. Wachinski, A. M. *Environmental Ion Exchange : Principles and Design, Second Edition*. (CRC Press, 2016). doi:10.1201/9781315368542
7. Clemens, S., Palmgren, M. G. & Krämer, U. A long way ahead: understanding and engineering plant metal accumulation. *Trends Plant Sci.* **7**, 309–315 (2002).
8. Cobbett, C. & Goldsbrough, P. Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. *Annu. Rev. Plant Biol.* **53**, 159–182 (2002).
9. Sousa, C., Cebolla, A. & de Lorenzo, V. Enhanced metalloadsorption of bacterial cells displaying poly-His peptides. *Nat. Biotechnol.* **14**, 1017–1020 (1996).
10. Bae, W., Wu, C. H., Kostal, J., Mulchandani, A. & Chen, W. Enhanced Mercury Biosorption by Bacterial Cells with Surface-Displayed merr. *Appl Env. Microbiol* **69**, 3176–3180 (2003).

11. Kuroda, K. & Ueda, M. Bioadsorption of cadmium ion by cell surface-engineered yeasts displaying metallothionein and hexa-His. *Appl. Microbiol. Biotechnol.* **63**, 182–186 (2003).

12. Pazirandeh, M., Wells, B. M. & Ryan, R. L. Development of bacterium-based heavy metal biosorbents: enhanced uptake of cadmium and mercury by *Escherichia coli* expressing a metal binding motif. *Appl Env. Microbiol* **64**, 4068–4072 (1998).

13. Ruta, L. L. *et al.* Heavy metal accumulation by *Saccharomyces cerevisiae* cells armed with metal binding hexapeptides targeted to the inner face of the plasma membrane. *Appl. Microbiol. Biotechnol.* **101**, 5749–5763 (2017).

14. Waldron, K. J. & Robinson, N. J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol.* **7**, 25–35 (2009).

15. P. Stathi, I.T. Papadas, A. Tselepidou & Y. Deligiannakis. Heavy-metal uptake by a high cation-exchange-capacity montmorillonite: the role of permanent charge sites. *Glob. NEST J.* **12**, 248–255 (2013).

16. Barakat, M. A. New trends in removing heavy metals from industrial wastewater. *Arab. J. Chem.* **4**, 361–377 (2011).

17. Rutherford, N. & Mourez, M. Surface display of proteins by Gram-negative bacterial autotransporters. *Microb. Cell Factories* **5**, 22 (2006).

18. Kieke, M. C. *et al.* Selection of functional T cell receptor mutants from a yeast surface-display library. *Proc. Natl. Acad. Sci.* **96**, 5651–5656 (1999).

19. Robertson, J. G. Determination of subunit dissociation constants in native and inactivated CTP synthetase by sedimentation equilibrium. *Biochemistry* **34**, 7533–7541 (1995).

20. Ingerson-Mahar, M., Briegel, A., Werner, J. N., Jensen, G. J. & Gitai, Z. The metabolic enzyme CTP synthase forms cytoskeletal filaments. *Nat. Cell Biol.* **12**, 739–746 (2010).

21. Liu, J.-L. The enigmatic cytoophidium: Compartmentation of CTP synthase via filament formation. *BioEssays* **33**, 159–164 (2011).

22. Carcamo, W. C. *et al.* Induction of cytoplasmic rods and rings structures by inhibition of the CTP and GTP synthetic pathway in mammalian cells. *PLOS ONE* **6**, e29690 (2011).

23. Liu, J.-L. Intracellular compartmentation of CTP synthase in *Drosophila*. *J. Genet. Genomics* **37**, 281–296 (2010).

24. Narayanaswamy, R. *et al.* Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation. *Proc. Natl. Acad. Sci.* **106**, 10147–10152 (2009).

25. Noree, C., Sato, B. K., Broyer, R. M. & Wilhelm, J. E. Identification of novel filament-forming proteins in *Saccharomyces cerevisiae* and *Drosophila melanogaster*. *J. Cell Biol.* **190**, 541–551 (2010).

26. Eisenberg, D., Gill, H. S., Pfluegl, G. M. U. & Rotstein, S. H. Structure–function relationships of glutamine synthetases. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1477**, 122–145 (2000).

27. Valentine, R. C., Shapiro, B. M. & Stadtman, E. R. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from *Escherichia coli*. *Biochemistry* **7**, 2143–2152 (1968).

28. Miller, R. E., Shelton, E. & Stadtman, E. R. Zinc-induced paracrystalline aggregation of glutamine synthetase. *Arch. Biochem. Biophys.* **163**, 155–171 (1974).

29. Robertson, J. G. & Villafranca, J. J. Characterization of metal ion activation and inhibition of CTP synthetase. *Biochemistry* **32**, 3769–3777 (1993).

30. Denton, M. D. & Ginsburg, A. Conformational changes in glutamine synthetase from *Escherichia coli*. I. Binding of manganese ion in relation to some aspects of the enzyme structure and activity. *Biochemistry* **8**, 1714–1725 (1969).

31. Knecht, S., Ricklin, D., Eberle, A. N. & Ernst, B. Oligohis-tags: mechanisms of binding to Ni^{2+}-NTA surfaces. *J. Mol. Recognit.* **22**, 270–279 (2009).

32. Greenleaf, J. E., Lin, J. & Sengupta, A. K. Two novel applications of ion exchange fibers: arsenic removal and chemical-free softening of hard water. *Environ. Prog.* **25**, 300–311 (2006).

33. Haiech, J., Klee, C. B., Demaille, J. G. & Haiech, J. Effects of cations on affinity of calmodulin for calcium: ordered binding of calcium ions allows the specific activation of calmodulin-stimulated enzymes. Theoretical approach to study of multiple ligand binding to a macromolecule. *Biochemistry* **20**, 3890–3897 (1981).

34. Kelley, B. Industrialization of mAb production technology. The bioprocessing industry at a crossroads. *mAbs* **1**, 443–452 (2009).

Acknowledgements

We would like to thank the Koch Institute Facility, especially Dr. Dong Soo Yun from the Nanotechnology Code for helping with HRTEM imaging. We would also like to thank the Center for Material Science Facilities (CMSE), especially Dr. Yong Zhang for TEM guidance and help. This work was supported by the Amar G. Bose Research Grant (G.L.S. and A.M.B.) and the NSF Graduate Fellowship (G.L.S.). Partial support for this research was provided by a core center grant P30-ES002109 from the National Institute of Environmental Health Sciences, National Institutes of Health.

Author contributions

G.L.S. and A.M.B. conceived the study and designed experiments; G.L.S. performed the experiments; G.L.S. analyzed the data and assembled figures; G.L.S. and A.M.B wrote the manuscript.

Competing interests

The authors declare no competing interests.