

# Chapter 1

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

### Abstract

Physicochemical technologies have dominated water treatment methods; however, impact on reducing contaminated waters has lagged behind the growing accumulation of heavy metal waste. Ion-exchange is one such method which has gained traction in developed countries such as the United States, but remains difficult to deploy in developing nations which are in more need of economical and sustainable remediation technologies. Unfortunately, ion-exchange requires sophisticated design strategies, advanced chemical synthesis techniques, and overall has a high cost and technical barrier for adoption. But with a closer look, the main mechanism of ion-exchange is straightforward and operates through reversible metal chelation, a reaction found almost everywhere in biological systems such as proteins and enzymes. 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

metal-inducible polymerization behavior in which they aggregate into solid matrixes when exposed to metals. Their surfaces can be decorated with either a 6xHis tag, metallothionein, or calmodulin fusions to tailor removal of copper, cadmium, mercury, lead, and even calcium. These proteins were further modified to harbor binding tags to allow anchorage onto denser substrates such as agarose, magnetic beads, or yeast to increase metal-protein sedimentation rates. With this system, modified pyrG and glnA showed more than 50% removal of 1 millimolar amounts of metal (more than 10–100 times higher than EPA thresholds for dangerous metal content) and more than 80% recovery of metals removed. The ability for proteins to act as ion-exchange like agents opens the potential for robust heavy metal removal and possible downstream collection and/or recycling.

## 1.1 Introduction

One of the most common methods for wastewater remediation is through ion-exchange. Ion-exchange is a well-established technology with high removal efficiency and fast removal kinetics [1–4]. When broken down to its fundamental mechanism, ion-exchange uses beads (also known as resins) that are functionalized with multiple copies of commonly used metal binding functional groups (e.g. sulfonic,  $-\text{SO}_3\text{H}$ , and carboxylic acids,  $-\text{COOH}$ ) [1]. Unfortunately, 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 required to aid in the design strategy and synthesis process [6]. What this design flow suggest is creating, testing, and manufacturing resins is arduous and requires technical skill. More so, during synthesis many of the chemical processes are prone to generating secondary-waste due to reaction by-products or effluent from regeneration of resins when preparing or after use [4, 5].

Looking again at the fundamental components of ion-exchange there are two main groups to engineer, the resin, and the identity and composition of 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 binders perform 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 serve no remediation function. But both components do not have to be synthetic, rather, natural biological systems have already developed a mirror-image approach for handling heavy metal waste. One such example is the removal of metals through chelation by polymerized chains of phytochelatins, or binding onto metallothionein proteins [7, 8]. In addition, the functional groups of strong and weak exchangers (e.g.  $-S^{2-}$ ,  $-NH_2$ , and  $-COOH^-$ ) are regularly found, and often modifiable, in most peptides/proteins. This approach has been recognized by scientists in the past, as some have attempted to mimic ion-exchange by functionalizing cell surfaces with metal binding peptides and proteins. The cell acts as the physical anchor 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 metallothioneins [10–12]. Compared to synthetic manufacturing of 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 domain [13], the diversity and richness of 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 water remediation applications.

However, past results in biologically mimicked ion-exchange is not yet able to compete with synthetic ion-exchange with respects to metal removal capacities. Results for cell surface display metal removal have capture capacities of tens of nanomolar of metal per gram of cell dry weight [9, 11, 14], whereas ion-exchange have typical capture capacities of millimolar of metal per gram dry weight, almost 3–6 orders of

magnitude more [15, 16]. Being that metal removal is purely a numbers game, the limitation in biologically mimicked ion-exchange is the low number of peptide/protein expression per cell per volume. For example, typical values of protein expression for yeast display range from 1 to 100 thousand [17, 18]. Given modest estimates of yeast culture density (assuming 1 OD<sub>600</sub> per mL) an upper bound of metal removal assuming each protein has 1 completely occupied binding site is roughly 2–200 nanomolar, an almost insignificant amount when dealing with micro to millimolar levels of metal contaminants.

In order to capitalize on the benefits of using biology, e.g. diversity, engineerability, scale, this work focuses on using proteins recently discovered to have polymerization properties for heavy metal chelation and sedimentation. Using proteins that can polymerize increase protein quantity per volume, and relying on proteins rather than an entire cell for metal capture would increase capture capacities per dry weight. Furthermore, proteins could be modified using facile genetic and protein engineering techniques to express metal binding moieties on their surfaces. Thus when polymerized the supramolecular complex can be decorated with multiple copies of metal chelators that are then pulled down during aggregation. Ultimately the proteins themselves, once aggregated, act as the resin and the fused metal binding moieties act as the exchangers. The two proteins investigated herein are bacterial pyrG from the CTP synthase (CS) 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 CS was observed to form cytoplasmic filaments in a conserved fashion among a variety of organisms [19–23]. This peculiar behavior has been associated with subcellular protein organization for enzyme regulation [20, 22, 23]. Given these observations, a multitude of other 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 interesting protein family, GS, was determined to also form supramolecular complexes, and some of the first recorded observations were cited as early as the 1970s [25–28]. In the context of metal removal, the particular enzymatic function, and even the cellular role of these

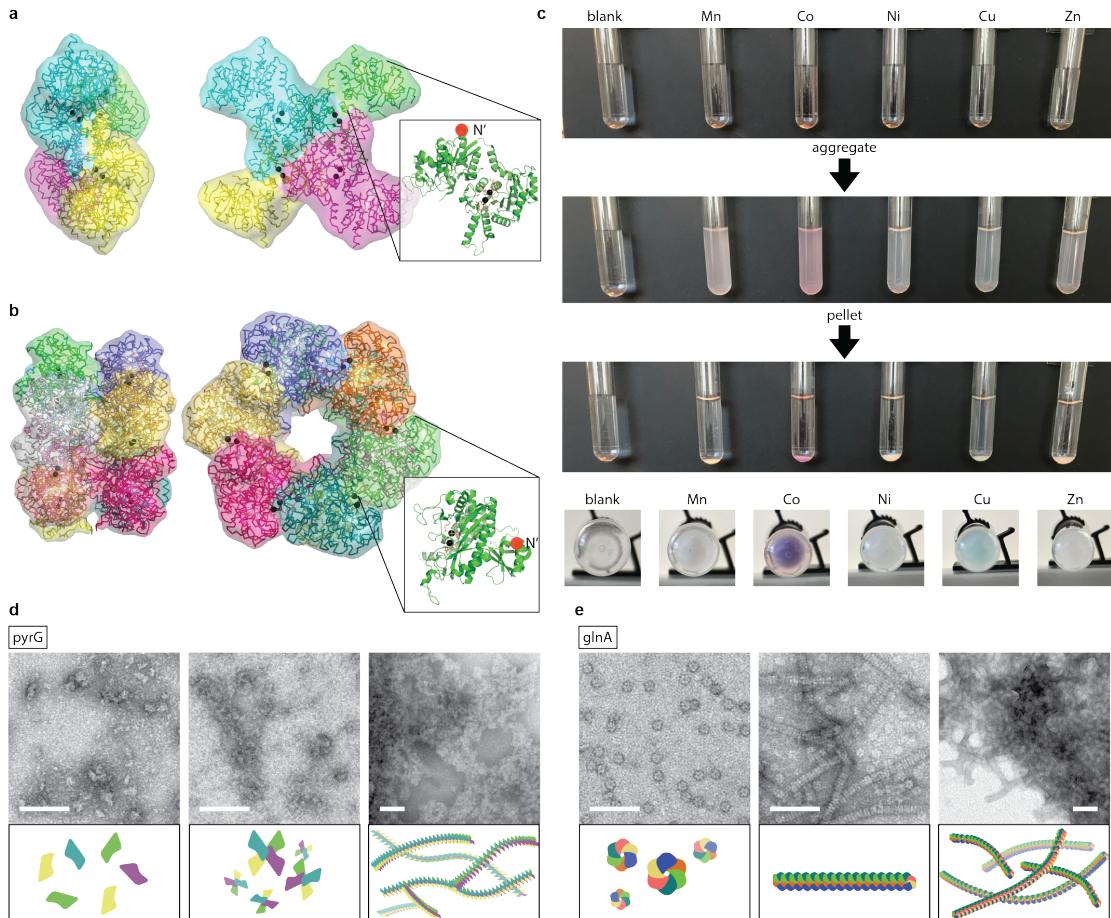
proteins, is irrelevant; the main interest is the fact that both CS and GS have controllable means to initiate polymerization and metal binding, and the focus of this work is aimed to capitalize on this mechanisms for a direct application in heavy metal waste removal.

## 1.2 Results

### 1.2.1 Using CS and GS for toxic metal removal

Representative crystal structures for pyrG (accession number #P0A7E5) and glnA (#P0A9C5) (PDB 1S1M and 1FPY, respectively) reveal that these proteins are monomers of a larger symmetric structure (Figure 1.1a,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 DH5 $\alpha$  *E. coli* strain and inserted into IPTG inducible pET28c(+) vector. Confirmed sequences were transformed into BL21(DE3) high expression strain, induced, lysed, and purified with Ni-NTA columns (Supplemental Figure 1.1) and protein concentrations quantified using calibrated Nanodrop readings (Supplemental Figure 1.2). Final protein samples and experimental reactions were performed in water. Past studies have shown that glnA aggregates in the presence of Mn(II) [27, 28], however a surprising observation is that both pyrG and glnA aggregate in response to a variety of other divalent metals such as Co(II), Ni(II), Cu(II), and Zn(II) (future reference to metal elements are assumed to be divalent from hereon; Figure 1.1c; Supplemental Figure 1.3). At 100  $\mu$ M pyrG or glnA added to 1 mM of Mn, Co, Ni, Cu, Zn, Cd, Hg, or Pb visibly turned cloudy and eventually formed aggregates which could be pelleted. To test whether 1 mM of metal was causing protein denaturation or non-specific aggregation, a control of 100  $\mu$ M BSA was added to 1 mM of the same metals, and no visible opaqueness formed, except for Pb which naturally formed hydroxides over time (Supplemental Figure 1.3). Alkaline and alkaline earth metals (Na, Mg, etc.) were not observed to

induce aggregation.



**Figure 1.1 | Using *pyrG* and *glnA* as aggregating metal binders 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 modeled 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  $\mu$ M respectively. (d) 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.

The aggregating tendencies of *pyrG* can be observed visually; *pyrG* forms spontaneous aggregates at room temperature, and further examination under transmission

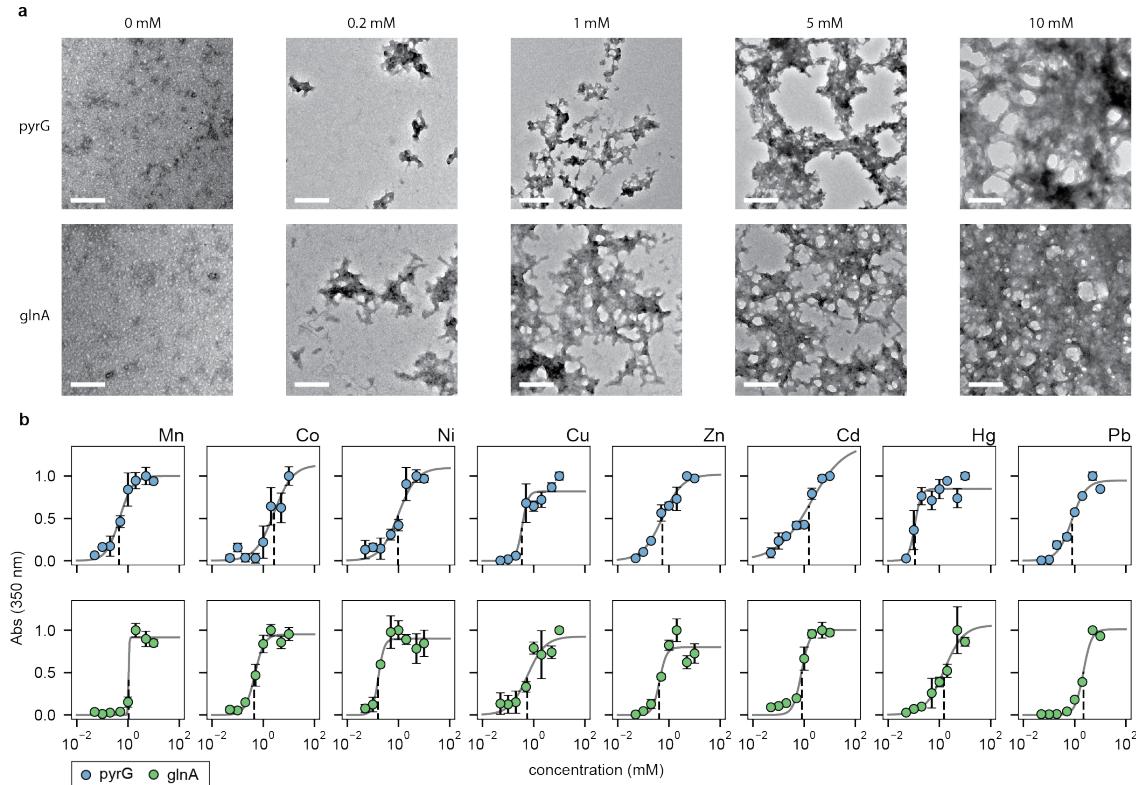
electron microscopy (TEM) show haphazard connections rather than its canonical tetrameric structure. Under high resolution TEM (HRTEM), pyrG was observed to not regularly form its tetrameric structures, but instead bundle as monomers. In the presence of metals pyrG forms a combination of tetrameric but more often irregular aggregated structures, and at the extreme aggregates into amorphous supramolecular complexes (Figure 1.1d; Supplemental Figure 1.4a). On the other hand, purified glnA examined under HRTEM forms uniform dodecamer structures. Upon addition of metal these structures aggregate into rods, with each hexagon 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 1.1e; Supplemental Figure 1.4b). This study utilizes the metal responsiveness of pyrG and glnA to chelate heavy metals away from waters for bioremediation applications.

### 1.2.2 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 induce 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 [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, a wavelength that produced the highest signal to noise ratio of aggregates to non-aggregated protein and showed a linear dependence on aggregation intensity in solution (Supplemental Figure 1.5). 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 increased with increasing metal concentrations (Figure 1.2a). The study was performed again by quantifying aggregates at 350 nm to model a binding (in this case, aggregation) curve. Ab-

sorbance 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 1.1)



**Figure 1.2 | Measuring metal induced aggregation responsiveness and intensity for pyrG and glnA.** (a) Top/bottom rows are pyrG and glnA respectively. Columns represent concentrations of Zn which was used to induce aggregation, with progressively higher concentrations leading to higher levels of aggregation. Scale bars represent 200 nm for all images. (b) pyrG and glnA were titrated with metals to measure the level of aggregation as a function of metal concentration. Aggregation was quantitatively measured using 350 absorbance readings (Supplemental Figure 1.5). For all data, the mean  $\pm$  s.d. of three replicates are shown.

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 >

$\text{Co} > \text{Zn} > \text{Cu} > \text{Cd} > \text{Mn} > \text{Hg} > \text{Pb}$  (Figure 1.2b; Table 1.1). The Hill coefficient for pyrG and glnA for most metals exceeded 1, suggesting positive cooperativity for aggregate formation.

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 (Supplemental Figure 1.6a). 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 (Supplemental Figure 1.6b). 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.

	Mn	Co	Ni	Cu	Zn	Cd	Hg	Pb	
A	$0.10 \pm 0.01$	$0.15 \pm 0.07$	$0.20 \pm 0.08$	$0.37 \pm 0.04$	$0.20 \pm 0.01$	$0.39 \pm 0.05$	$0.24 \pm 0.01$	$0.38 \pm 0.01$	
pyrG	$K_D$	$0.5 \pm 0.01$	$4.66 \pm 0.43$	$3.7 \pm 0.4$	$0.42 \pm 0.2$	$0.57 \pm 0.03$	$2.44 \pm .12$	$0.13 \pm 0.05$	$0.76 \pm 0.04$
n		$1.58 \pm 0.21$	$12.43 \pm 1.6$	$1.67 \pm 0.87$	$11.06 \pm 11.4$	$1.1 \pm 0.06$	$0.66 \pm 0.15$	$4.31 \pm 0.20$	$1.59 \pm 0.22$
A	$0.22 \pm 0.01$	$0.15 \pm 0.01$	$0.12 \pm 0.01$	$0.21 \pm 0.03$	$0.16 \pm 0.02$	$0.18 \pm 0.01$	$0.08 \pm 0.01$	$1.37 \pm 0.01$	
glnA	$K_D$	$1.05 \pm 0.02$	$0.48 \pm 0.06$	$0.17 \pm 0.01$	$0.61 \pm 0.14$	$0.42 \pm 0.05$	$0.8 \pm 0.09$	$2.11 \pm 0.86$	$2.18 \pm 0.04$
n		$32.86 \pm 3.15$	$3.32 \pm .18$	$4.57 \pm .18$	$9.93 \pm 1.3$	$3.25 \pm 0.5$	$2.96 \pm 0.76$	$1.33 \pm 0.6$	$2.7 \pm 0.14$

**Table 1.1 | Values for maximum aggregation intensity ( $A$ ; measured at 350 nm),  $K_D$ , and fitted cooperativity coefficient ( $n$ ) for pyrG and glnA for the various metals studied.** Coefficients were fitted from data collected in Figure 1.2b.

### 1.2.3 Metal sequestration via protein-metal aggregation

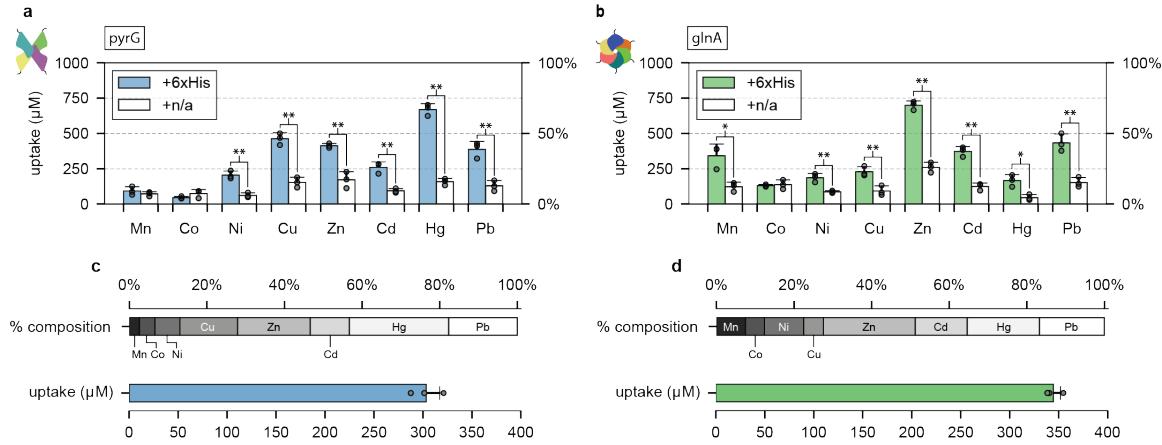
Beyond understanding the aggregating behaviors of pyrG and glnA, the main objective of this work was to chelate and sediment metals upon induced aggregation. To increase binding capacities per pyrG and glnA complex, their surfaces were decorated with metal binding domains. The exposed N'-terminus was used to fuse several metal

binding domains (Figure 1.1a,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  $\mu$ M of protein and 1 mM metal for aggregation and metal removal studies.

Two mutually dependent factors contribute to the level of metal removal. The first is the degree of metal-induced aggregation which physically assembles the protein complex and allow it to sediment out of 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, irregardless of the strength of metal binding the protein-metal complex will remain in solution if it does not aggregate and sediment. On the other hand, if there was no metal binding 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 1.3a,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 removal when compared to glnA ( $p < .05$ ). Compared to controls of 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 ( $463 \pm 34 \mu\text{M}$ ), Zn ( $413 \pm 13 \mu\text{M}$ ), and Hg ( $670 \pm 33 \mu\text{M}$ ) in the presence of 1 mM metal. Whereas for glnA were with Zn ( $699 \pm 25 \mu\text{M}$ ), Cd ( $373 \pm 28 \mu\text{M}$ ), and Pb ( $433 \pm 51 \mu\text{M}$ ) 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.

pyrG and glnA were also tested in mixed metal conditions. Mn, Co, Ni, Cu, Zn, Cd, Hg, and Pb were mixed at equal molar to a final metal concentration of 1 mM ( $125 \mu\text{M}$  per metal) and mixed with 100  $\mu\text{M}$  pyrG and glnA. The amount of metal removed for both were approximately  $>33\%$  (Figure 1.3c,d). For pyrG, the

individual metal makeup of what was removed favored Cu, Zn, Cd, Hg, and Pb. A similar analysis for glnA showed a similar composition of metal removed from the mixed source (Figure 1.3c,d; Supplemental Figure 1.7).



**Figure 1.3 | Metal removal of Mn, Co, Ni, Cu, Zn, Cd, Hg, and Pb for pyrG and glnA modified with a 6xHis tag.** (a, b) Removal of added 1 mM metals were individually measured for pyrG or glnA, respectively. Controls (white bars) represent pyrG and glnA cleaved of its 6xHis tag fusion and measured for metal removal. (c, d) Metal removal of a metal mixture (totaling 1 mM, or 125  $\mu\text{M}$  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. \* =  $p < .05$ , and \*\* =  $p < .01$  represent significant metal removal compared to controls.

Even though pyrG and glnA have different metal aggregating responses, the presence of all 8 metals strongly encouraged pyrG and glnA aggregation as the weakest and strongest metal-inducers were present. Therefore aggregation was no longer an issue and the responsibility for metal removal was now determined by the 6xHis tag. Therefore, the metal removal composition profile appear roughly the same for pyrG and glnA because of the same 6xHis metal-binding domain. To better illustrate this takeaway: glnA has a relatively weak aggregating effect with Hg, despite the 6xHis tag having a relatively strong affinity for Hg (Figure 1.3b). 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. 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 (Supplemental Figure 1.8). The results shown in Figure 1.3c,d is an extension of this example, where the composition of metal removed is more dependent on the 6xHis tag.

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 of metal removal. pyrG and glnA have two native metal binding sites per monomer (Figure 1.1a,b), and the addition of the 6xHis tag now increases the stoichiometry to at least 3. However, the calculated stoichiometry of pyrG and glnA given the metal removal results (Figure 1.3) give ratios between 3–8 (depending on the metal), higher than expected. Several explanations can be made: the 6xHis tag could bind to multiple metals, 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 at least 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 may positively encourage metal binding to a greater amount than compared to the predicted one 6xHis tag to one metal assumption. Therefore, modifications to the protein surface is an important engineering parameter to design for in order to both enhance metal removal specificity and metal removal capacity.

#### **1.2.4 Tuning metal removal profiles by selecting new metal-binding domains**

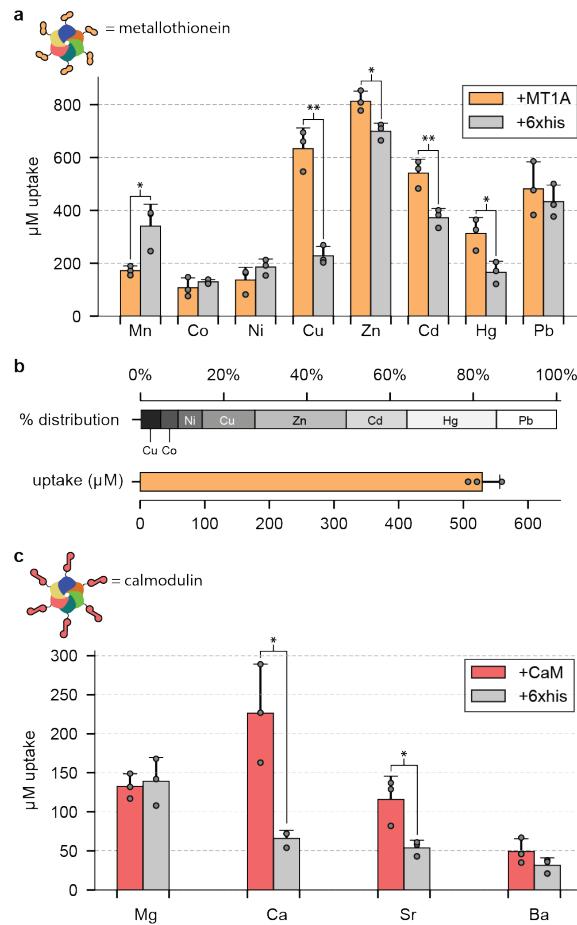
Aside from the 6xHis tag, there are many other metal-binding peptides/proteins that can serve as metal-binding accomplices with pyrG or glnA. The remainder of this study focuses on glnA because of its better stability, more defined macrostructure as examined under TEM, and the higher metal-protein binding stoichiometry per

macromolecule compared to pyrG (12 compared to 4).

A family of metal-binding proteins used in a variety of cells for metal detoxification are metallothioneins (MTs). MTs are cysteine-rich, low molecular weight proteins with high metal affinity for common toxic metals such as copper, cadmium, and mercury [8]. Many of the isoforms of MTs have been studied in plants, and because *E. coli* do not natively have MTs, the MT1A (#P43392) from *A. thaliana* was codon-optimized and fused to the N'-terminus of glnA which also replaced the 6xHis tag. Performing the identical metal removal experiment as before, the profile of metals removed skewed more towards Cu, Zn, Cd, Hg. The removal of Cu, Zn, Cd, and Hg increased by  $406 \pm 92 \mu\text{M}$ ,  $110 \pm 13 \mu\text{M}$ ,  $169 \pm 62 \mu\text{M}$ , and  $147 \pm 53 \mu\text{M}$ , respectively, in comparison to using the 6xHis tag (Figure 1.4a). In a multi-metal removal experiment the enhanced metal removal is more apparent, with total metal removal exceeding 50% and a large composition of the metal removed being Cu, Zn, Cd and Hg (Figure 1.4b).

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 decrease remediation efficiency processes like ion-exchange [32]. A commonly known and used protein for calcium binding, calmodulin (CaM), was used to test whether the metal removal profile could be shifted towards the alkaline-earth metals. Yeast CaM (#P06787) was isolated from the genome of *S. Cerevisiae* W303 $\alpha$  strains and cloned into glnA to replace the N'-terminus 6xHis tag. 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  $226 \pm 51 \mu\text{M}$  compared to  $66 \pm 8 \mu\text{M}$  of glnA with 6xHis ( $p < .01$ ), and a slight increase in Sr removal of  $116 \pm 30 \mu\text{M}$  compared to  $54 \pm 7.8 \mu\text{M}$  of glnA with 6xHis ( $p < .05$ ) (Figure 1.4c). The metal removal profile loosely

corresponded to CaM native preference for Ca, followed by Mg and Sr [33].

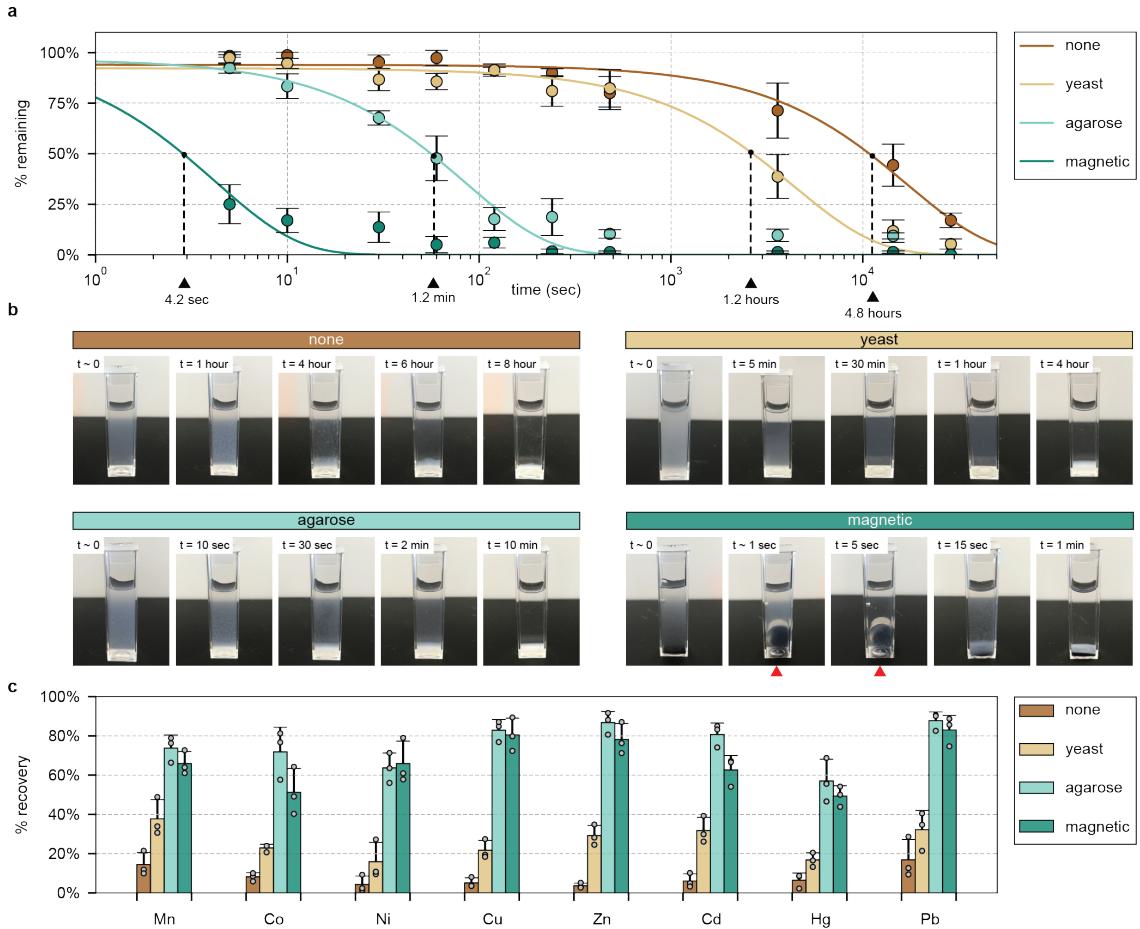


**Figure 1.4 | Substituting the 6xHis tag with plant MT1A or yeast calmodulin alters metal binding preference and metal removal capacity.** (a) The same metal removal experiment was performed for glnA+MT1 and compared against glnA+6xHis. Individual metal removal preference skewed more towards Cu, Zn, Cd, and Hg ( $p < .05$ ). (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) glnA+CaM were aggregated with Zn and tested for Mg, Ca, Sr, and Ba removal. For all data, the mean  $\pm$  s.d. of three replicates are shown. \* =  $p < .05$ , and \*\* =  $p < .01$  represent significant metal removal compared to controls.

### 1.2.5 Improving sedimentation of 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) (Supplemental Figure 1.9). glnA aggregation was performed together with yeast displaying cells such that 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  $\mu$ M 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 1.5a). 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 1.5a,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.



**Figure 1.5 | glnA fused with a 3XFlag tag, or bound to yeast displaying glnA monomers, improved 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. Alternatively, 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 protein-metal complex sedimentation with magnetic, agarose beads, and yeast. Red arrows for the magnetic bead condition indicate moments when an external magnet was applied. (c) Mixtures were allowed to sediment for 10 minutes in which the sedimented pellet was isolated, washed, and extracted for 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  $\pm$  s.d. of three replicates are shown.

Metal recovery was also measured using the various sedimentation strategies. To account for the maximum amount of metal recoverable, the amount of metal bound to the glnA complex was first measured. Afterwards, complexes were then resuspended

and mixed with agarose or magnetic beads. For the yeast display condition, glnA and yeast were mixed during metal induced aggregation in order to allow proteins to aggregate onto the yeast surface. After incubation, the protein-metal bead/yeast complex were stirred and allowed to sediment. Samples were allowed to sit for 10 minutes after which the supernatant was fully removed and whatever sedimented was recovered. The sediment was transferred and washed 2X in ddH<sub>2</sub>O before resuspending in 10 mM EDTA. The amount of freed metals were measured and compared against the initial measurement of metals bound to the aggregate before the sedimentation experiment. This value gave a metal percent recovery value in the 10 minute span allowed for sedimentation.

Factors that affected recovery were the rate of sedimentation within the 10 minute timeframe, and the binding strength between the aggregates and anchors during the wash steps. Without any anchor, the recovery for just glnA was below 10% for all metals (Figure 1.5c). For yeast it was observed that flakes of aggregates would dissociate during the wash steps; this was not seen for the agarose or magnetic bead samples. Aggregates with yeast displayed monomers had recoveries between 15–40%. When using beads, the agarose and magnetic anchors were quicker to sediment thereby recovering more aggregates which corresponded to a higher percent recovery of Cu, Zn, Cd and Pb. Metal recoveries ranged between 60–80% for the agarose and magnetic bead conditions. The high recoveries could also be associated to the tight binding of the 3XFlag tag between glnA and the functionalized antibodies on the beads. A takeaway from these experiments is that not only can glnA be modified to increase metal removal and metal removal preferences using appended metal binding domains, but further modifications can lead to other functionality such as attachment to beads or cells for improved sedimentation and recovery.

### 1.3 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 aggregating behavior and the subsequent solid sediment that forms allowing for physical handling and removal. 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 can be 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 native pyrG and glnA. 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 damage in water-related 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 could be 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 ca-

pacities and tailor metal removal preferences. For example, multiple repeats of the 6xHis tag can be fused to glnA 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 onto glnA 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 aggregating 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 overburdening the protein is true for most protein-fusion expression systems and requires fine tuning of expression conditions and purification protocol.

The identity of the protein anchor is not as important as its ability to aggregate and sediment out of solution. There has been much work in the biological community to identify the growing number of enzymes that behave like pyrG and glnA, and so far a list of 33 proteins discovered in yeast may have similar properties [24], and many more may exist in other organisms [22, 23]. Although not exhaustively tested here, there is a possibility that these identified proteins behave differently in either responsiveness or sensitivity towards aggregation, and these behaviors could be uniquely capitalized for different heavy metal removal conditions. More so, protein stability, stability with respects to surface modifications, and controlled aggregation are desired behaviors of the host protein. Future steps would be to understand mechanistically and biochemically how aggregation can be induced and controlled in this set of discovered proteins. If these underlying mechanisms can be uncovered, more engineering can be done to improve the aggregation sensitivity and density of aggregates

for improved metal removal, sedimentation, and recovery.

The goal of this work was to biologically mimic ion-exchange by recapitulating its underlying mechanism in proteins. The first was to use aggregating proteins to physically mimic the resin bed. The second was to fuse metal binding peptides/proteins onto these aggregates for targeted metal removal, a process similar to functionalizing resins with strong and weak metal exchangers. However, unlike ion-exchange these proteins are autonomously produced in cells and do not require complex chemical synthesis normally used when creating resins. In addition, 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 could 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 the others that many more sustainable solutions may exist with further exploration at the intersection of biology and technology.

## 1.4 Materials and methods

### Gene isolation and plasmid construction

Genomic DNA from *E. coli* DH5 $\alpha$  (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 (Supplemental Figure 1.2). 5' primer contained an overhang adding a TEV protease site (ENLYFQSQ) downstream of the pET28c(+) 6xHis tag before the gene insertion. Primers were used to amplify the appropriate genes using 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 NEB $\alpha$  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) upstream of the glnA sequence. The DNA sequence for CaM was isolated from the yeast genome and purified using the DNA Fungal/Bacterial Miniprep Kit (Zymo), whereas the MT1A sequence was codon-optimized and synthesized (GenScript). Primers were created for HiFi assembly into the pET28c(+-glnA vector (Supplemental Table 1.1) following the protocol already described above.

Alternatively, pET28c(+-glnA was modified by appending a 3XFlag tag upstream of glnA and downstream of the metal binding fusion (i.e. 6xHis tag). Primers were designed to anneal before the glnA sequence contained overhangs carrying the 3XFlag sequence (DYKDHDGDDYKDHDIDYKDDDDK) (Supplemental Table 1.2). The linear PCR product was re-circularized by adding T4 polynucleotide kinase with T4 ligase in 1X T4 buffer (NEB) and plasmid transformed in competent *E. coli*.

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 (Supplemental Table 1.3) and isolated using the protocol already described above. The pYD1 vector was linearized by cutting with NheI-HF and BamHI-HF restriction enzyme (NEB). pyrG or glnA were ligated into linearized pYD1 using T4 ligase and then transformed into competent *E. coli* and yeast W303 $\alpha$  strains.

## 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 30°C 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 -80°C 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  $\mu$ L of lysonase bioprocessing reagent (Millipore Sigma), 1  $\mu$ L of DNase (NEB), and 10  $\mu$ L 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 solution were spun down at 16,000xg for 20 minutes at 4°C 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 equilibration buffer removed. After bacterial lysis and centrifugation, supernatant was removed and filtered through a 0.45  $\mu$ m syringe filter directly onto equilibrated Ni-NTA resin. Mixtures were gently rocked for at least 2 hours at 4°C.

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 equates to 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 equates to 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 4°C. 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 stain. Protein samples were diluted to 0.5 mg in 20  $\mu$ L sample volume containing 1X LDS buffer and 1X denaturation buffer (Thermo). Samples were boiled at 70°C 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.  $[660] = m_{\text{nano}}[\text{Nano}] + b_{\text{nano}}$ . 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.  $[\text{mg/mL}] = m_{600}[660] + b_{660}$ . Substituting the relationship between the Nanodrop and the Pierce 660 nm protein readings, the final equation relating the Nanodrop to the BSA calibrated protein concentration reading of mg/mL gave  $[\text{mg/mL}] = (m_{660} \cdot m_{\text{nano}})[\text{Nano}] + m_{660} \cdot b_{\text{nano}}$ . 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 (Supplemental Figure 1.2).

Protein concentrations of mg/mL were converted to  $\mu\text{M}$  by using the monomer molecular weight. pyrG = 62.5 kDa, and glnA = 54 kDa.

## Aggregation quantification using absorbance measurement

pyrG and glnA without a 6xHis tag were used in absorbance reading 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 versus non-aggregated protein, hence highest signal-to-noise ratio (Supplemental Figure 1.2).

Proteins at 100  $\mu\text{M}$  were aggregated in 100  $\mu\text{L}$  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, intensity of aggregation ( $A$ ), and Hill coefficient ( $n$ ) (Table 1.1).

Testing reversibility of protein aggregation was performed by aggregating proteins at 1 mM metal for 1 hour. Resuspended aggregates were aliquoted in 100  $\mu\text{L}$  in a 96-well plate at a final concentration of 100  $\mu\text{M}$ . 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 removal experiments

Liquid stocks of manganese (II) chloride, cobalt (II) chloride, nickel (II) chloride, copper (II) chloride, zinc (II) chloride, cadmium (II) nitrate, mercury (II) chloride, and lead (II) nitrate (Sigma) were made at 100 mM in water. The same for the alkaline earth metals magnesium (II) chloride, calcium (II) chloride, strontium (II) chloride, and barium (II) chloride (Sigma) were made at 100 mM in water.

Protein samples were diluted to 100  $\mu$ M in water for metal removal experiments.

Protein aggregates were first formed by mixing 100  $\mu$ M 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°C. The supernatant was collected and diluted 1:10 in 3% HNO<sub>3</sub> solution for inductively coupled plasma (ICP) measurement. An ICP-OES (Agilent 5100) was used to measure metal concentrations of 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 formed 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 (i.e. 100  $\mu$ M) with the adjusted ICP measurement from the supernatant.

For multi-element removal 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  $\mu$ M). 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 removal experiments of control pyrG and glnA without a 6xHis tag were used. 6xHis tag of purified pyrG and glnA were cleaved using TEV protease, dialyzed, 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 experiments.

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  $\mu$ L of either Mg, Ca, Sr or Ba and 100  $\mu$ M 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  $\mu$ L 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  $\mu$ L of negative staining solution phosphotungstic acid (Sigma) at 1% was dropped on the grid and immediately removed with a kimwipe. Another 10  $\mu$ L 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.

## **Yeast display of CS and GS monomers**

Yeast EBY100 strain were made competent using Frozen-EZ Yeast Transformation II Kit (Zymo) and either stored at -80°C, 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  $\mu$ m glass beads (Sigma) and phenol:chloroform (Sigma) extraction and ethanol precipitation. A region of the pYD1 vector was amplified with primers flanking the AGA2 and T7 promoter using PCR and checked via gel electrophoresis for correct insertion.

Transformed strains were grown overnight in SDCAA media in 30°C, 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 4°C. Secondary antibodies conjugated with 488 or 647 dye were used against the primary antibodies and stained for 1 hour at 4°C. 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  $\mu$ M was aggregated with 1 mM of Cd in 5 mL for 1 hour. Samples were then mixed with 100  $\mu$ L of pure resin pre-equilibrated in ddH<sub>2</sub>O 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<sub>600</sub> of expressing EBY100 was mixed with 100  $\mu\text{M}$  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 fluorometer cuvette.

For sedimentation studies, samples were collected at the mid-height (e.g. 2 mL for a 4 mL cuvette). 20  $\mu\text{L}$  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  $\mu\text{L}$  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 (i.e. 100  $\mu\text{M}$ ) 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. So for agarose and magnetic beads, glnA without 3xFlag was used. For yeast display conditions 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; 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  $\mu\text{M}$  glnA+3xFlag and 1 mM metal were allowed to aggregate for 1 hour before spinning the complex 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 (i.e. maximum amount of metal recoverable, 100%). As a control, samples with no protein were measured to account for the natural precipitation of metal, namely Pb. This value was subtracted from the initial metal measurement to isolate the actual amount of metal removed from the protein-metal complex. After-

wards, aggregates were resuspended and then mixed with 100  $\mu\text{L}$  of pure agarose or magnetic resin pre-equilibrated in ddH<sub>2</sub>O. 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  $\mu\text{L}$  of sample was removed. The remaining volume was diluted with 1 mL of ddH<sub>2</sub>O, 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  $\mu\text{M}$  glnA was mixed with 1 mM metal and 1 OD<sub>600</sub> 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  $\mu\text{M}$  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

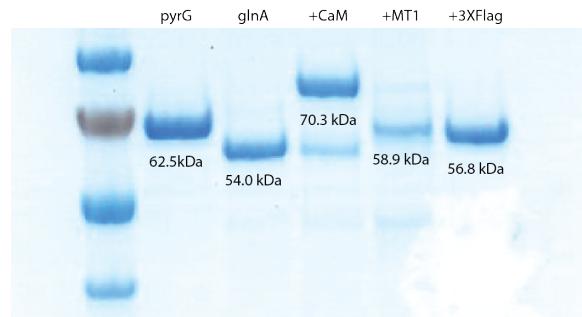
Raw data was collected and stored as csv or Excel file formats. Data was imported and analyzed with Python using modules such as numpy, pandas, and scipy. Plots

were graphed with matplotlib.

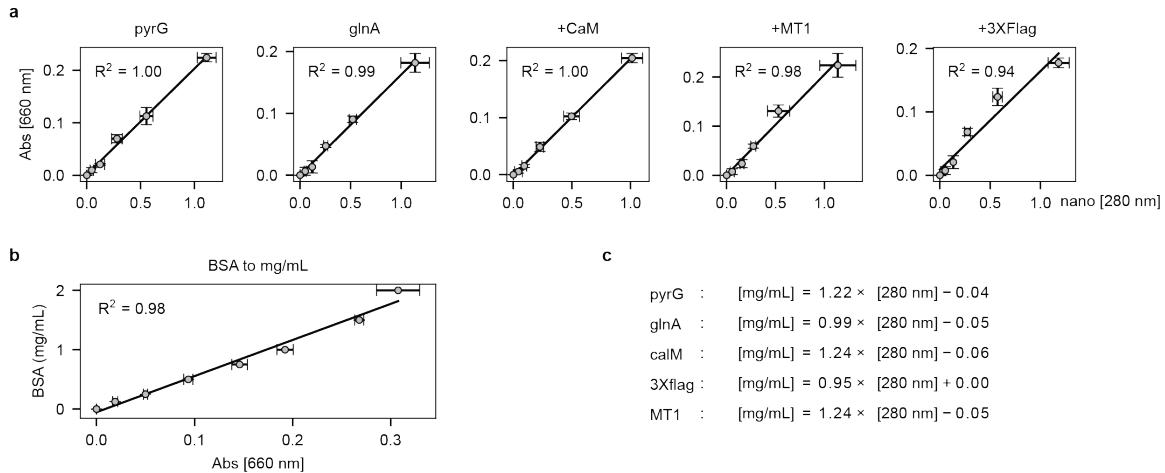
## Statistical analysis

Statistical parameters including the the definition 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 and p-values reported in the text. The significance threshold was set to  $p < .05$  for all experiments, or as specified in the text. In the figures, \* =  $p < .05$  and \*\* =  $p < .01$ .

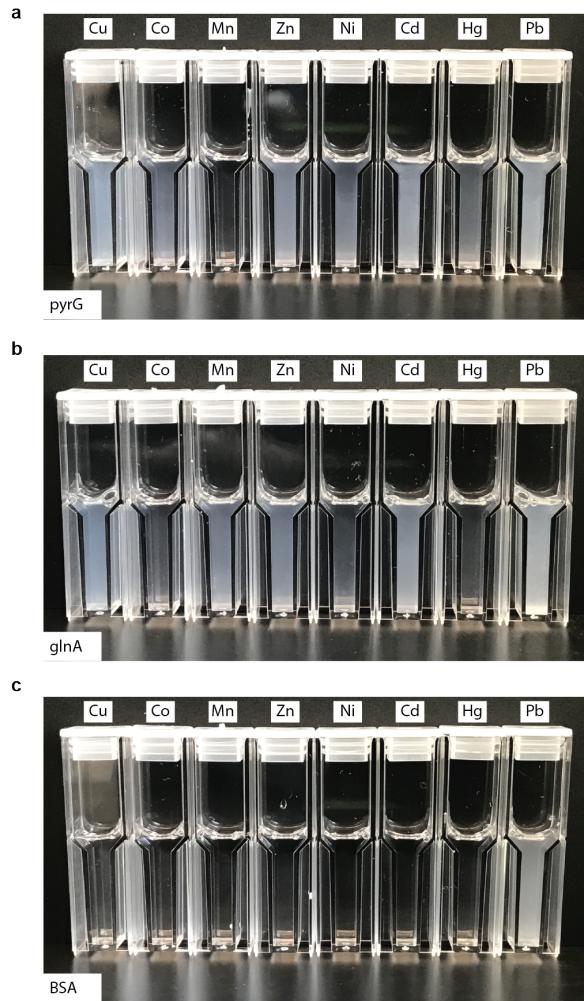
## 1.5 Supplemental figures



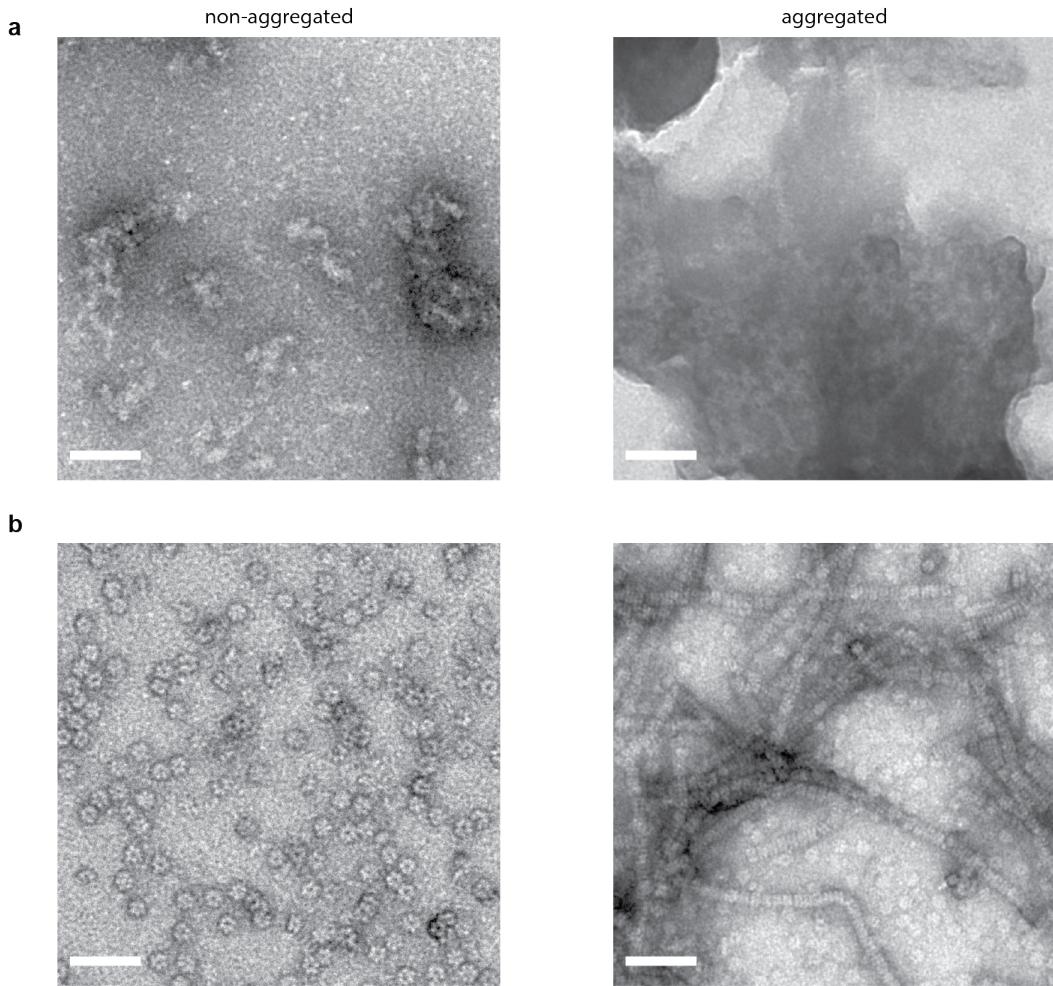
**S.Figure 1.1 | Examining protein expression and purity using SDS-PAGE and coomassie staining.** Purified pyrG, glnA, glnA+CaM, glnA+MT1, and glnA+3XFlag showed the correct band size. Only glnA+CaM showed a fainter second band which corresponds to un-modified glnA, suggesting that the glnA+CaM fusion was cleaved either during expression or during the purification process.



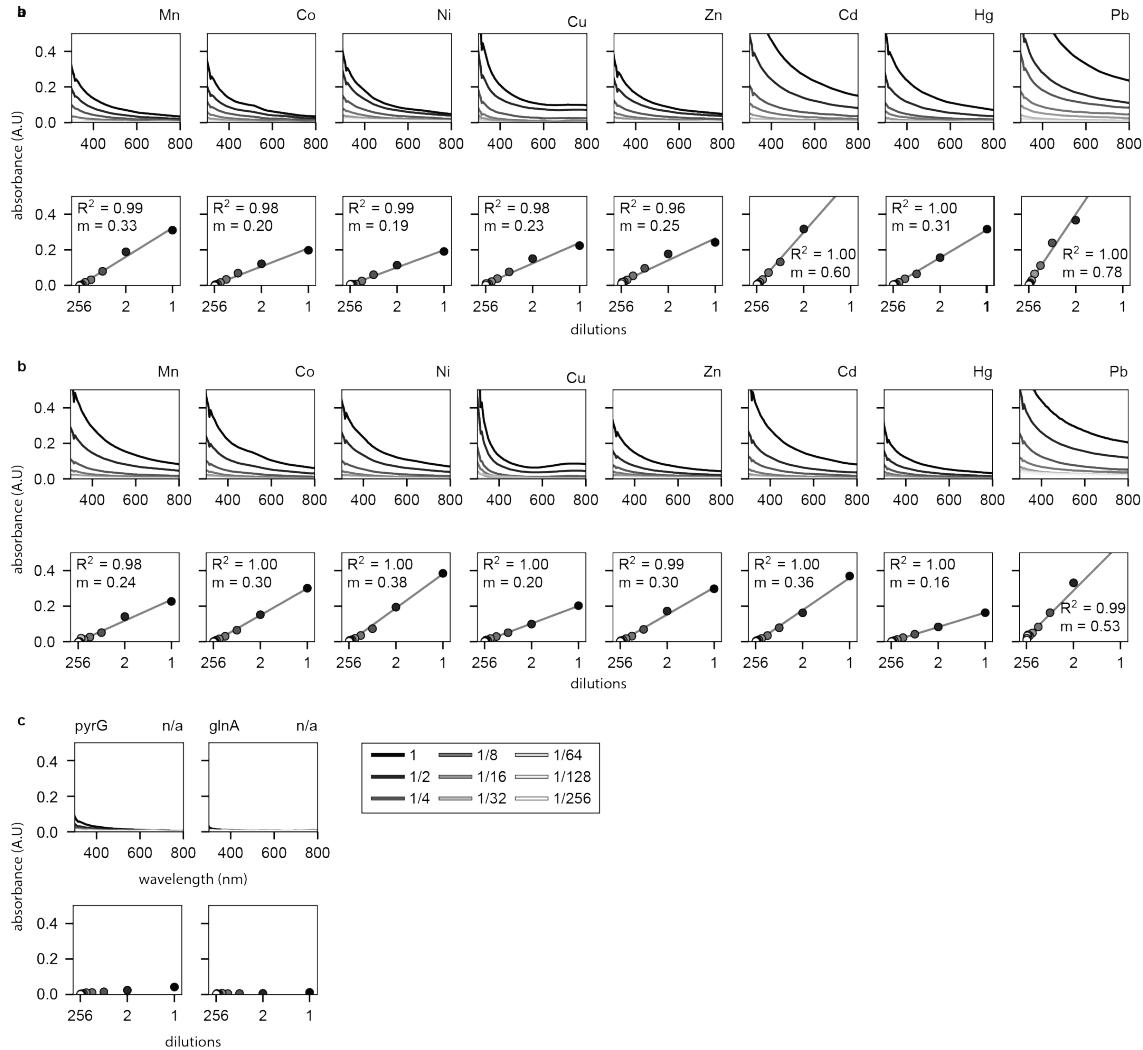
**S. Figure 1.2 | Protein concentrations measured on Nanodrop were correlated with Pierce 660 Protein Assay.** (a) Serial dilutions of proteins pyrG, glnA, +CaM, +MT1, and +3XFlag were measured on Nanodrop (x-axis) and Pierce 660 nm protein assay (y-axis). (b) Pierce assay was then used to create a calibration curve against BSA standards. The calibration curve, as well as the relationship between the Nanodrop and Pierce 660 nm protein assay was used to create a linear relationship (c) to better quantify concentrations for each protein. For all data, the mean  $\pm$  s.d. of three replicates are shown.



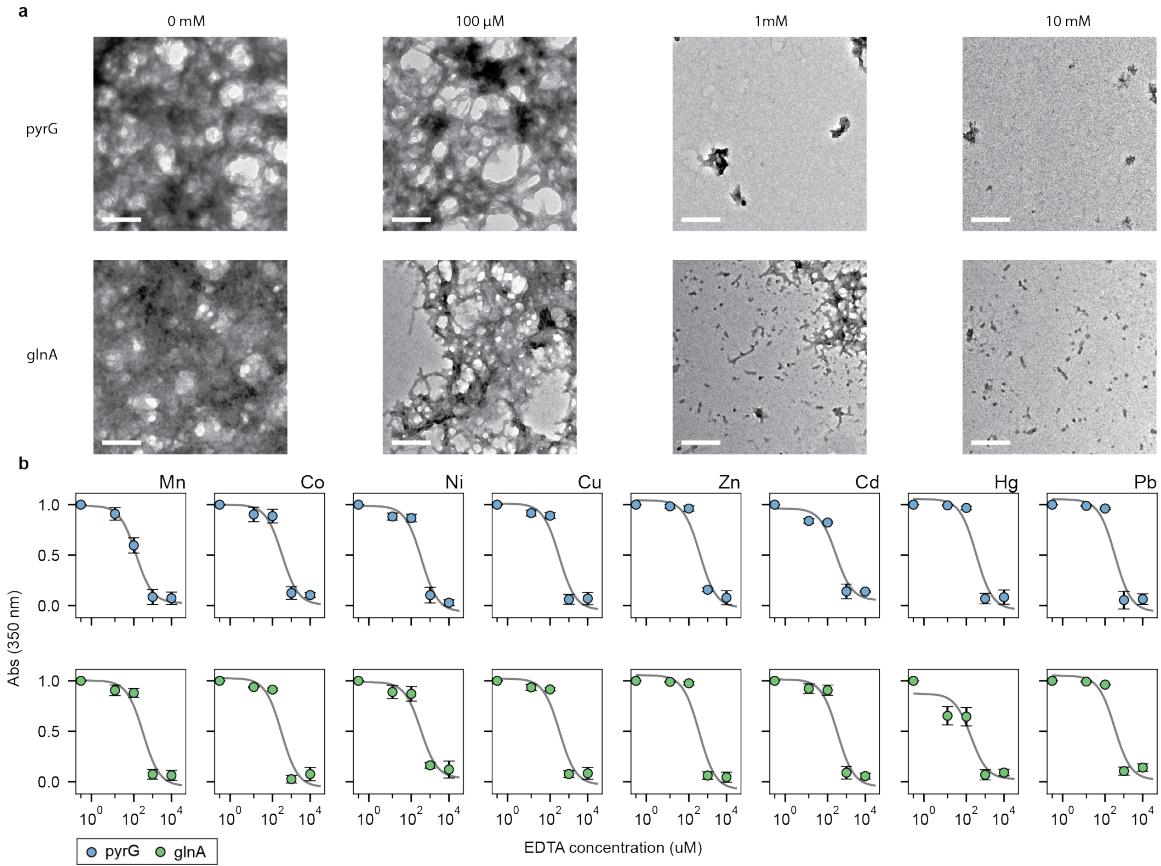
**S.Figure 1.3 | Images of 100  $\mu$ M pyrG, glnA and BSA mixed with 1 mM of metals.** (a, b) Aggregation of pyrG and glnA. (c) As a control, BSA was also mixed with 1 mM of metals to measure aggregation due to protein denaturation, or metal precipitation. The only visible change in opacity was with Pb, as it spontaneously formed hydroxides in neutral pH over time.



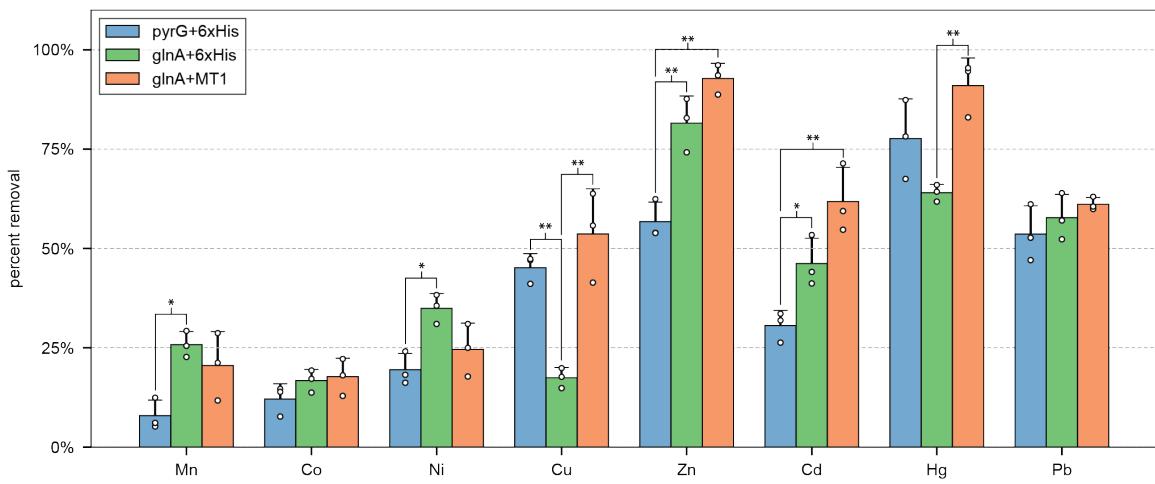
**S.Figure 1.4 | HRTEM images of pyrG and glnA.** (a) High resolution of pyrG show no definitive structure, but rather monomers that haphazardly form small aggregates. When aggregation is induced with metal, pyrG produces densely packed structures as seen by the dark contrast. (b) glnA forms the predicted structure of two hexamers stacked on top of one another. When aggregated, glnA forms a cob-like structures with glnA stacked on top of one another in a vertical fashion.



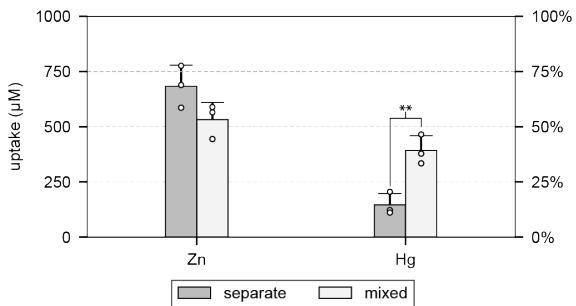
**S. Figure 1.5 | Absorbance scans of pyrG and glnA at varying degrees of metal-induced aggregation.** (a, b) pyrG and glnA absorbance scans for all metals tested in this study at increasing dilutions as indicated by the lower panel x-axis and increasing line opacity. Scans at 350 nm gave a linear relationship with good signal to noise for the various dilutions of protein aggregation. (c) Absorbance scans of non-aggregated pyrG and glnA absent of any metals.



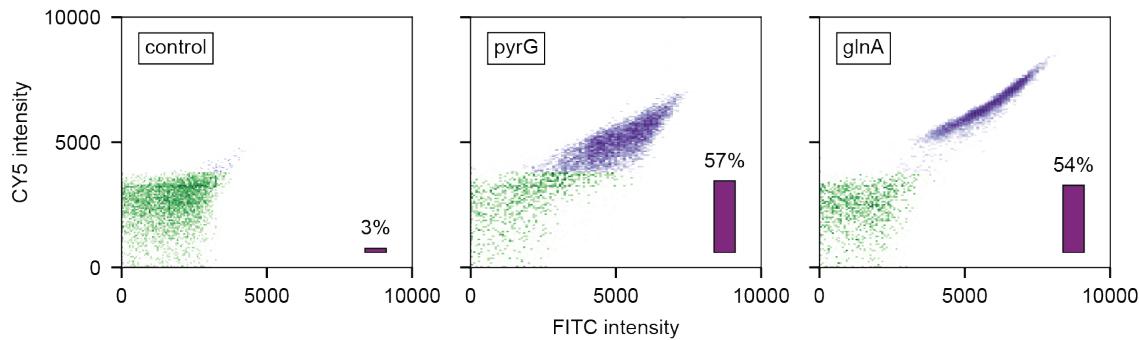
**S. Figure 1.6 | Reversibility of aggregated pyrG and glnA after metal removal using EDTA.** (a) Top/bottom row are pyrG and glnA, respectively. Columns represent concentrations of EDTA added after 100  $\mu$ M of protein was aggregated with 1 mM of Zn for 1 hour. Scale bars represent 200 nm for all images. (b) After protein aggregation, concentrations of EDTA (x-axis) were added to compete for metal chelation, and eventual dissociation of protein subunits. The intensity of aggregation was measured at 350 nm after adding EDTA for 10 min. For all data, the mean  $\pm$  s.d. of three replicates are shown.



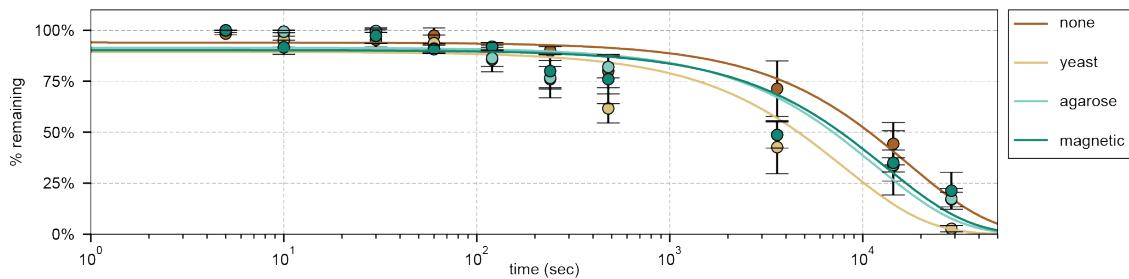
**S.Figure 1.7 | Percent metal removal for multi-metal removal experiments using *pyrG*, *glnA*, and *glnA* with MT1.** Multi-metal removal data reported in Figure 1.3c,d and Figure 1.4e were converted to percent fraction of metal removed. For all data, the mean  $\pm$  s.d. of three replicates are shown. \* =  $p < .05$ , and \*\* =  $p < .01$  represent significant metal removal compared to controls.



**S.Figure 1.8 | Mixing Zn and Hg alters metal removal for glnA.** Individually, Zn is removed significantly more than Hg when mixed with 100  $\mu\text{M}$  glnA with a 6xHis tag. However, when Zn and Hg are mixed together, Zn removal decreases while Hg increases ( $p < .01$ ). An explanation for this trend is that glnA does not aggregate as strongly with Hg it does with Zn. However, when mixed together, glnA aggregates in the presence of Zn but the 6xHis tag preferentially pulls down Hg over Zn. Therefore Hg removal piggy-backs Zn aggregation of glnA. For all data, the mean  $\pm$  s.d. of three replicates are shown. \*\* =  $p < .01$  represent significant metal removal compared to controls.



**S.Figure 1.9 | Yeast display of *pyrG* and *glnA* monomers using EBY100 and pYD1 vector.** Expression was analyzed using flow cytometry by staining the N'-terminus HA tag and C'-terminus Flag tag with antibodies conjugated with 488 (FITC) and 647 (Cy5) dyes. A WT control was measured in parallel to properly bin the population of non-expressing cells to expressing cells. The percent expression is shown as the lower right inlet bar chart.



**S.Figure 1.10 | Control sedimentation study of glnA using agarose or magnetic beads, or yeast.** glnA without a 3XFlag tag was aggregated and mixed with anti-flag agarose or magnetic beads. The yeast control had EBY100 strains transformed with an empty pYD1 vector and mixed with aggregated glnA. The sedimentation rates for each control was comparable to the sample containing no beads or yeast. For all data, the mean  $\pm$  s.d. of three replicates are shown.

## 1.6 Supplemental tables

---

name	direction	sequence
pyrG	fwd	CAGCAGCGCGAAAACCTGTATTTCAGAGCACAA CGAACTATATTTTGACCC
	rev	CCACCAGTCATGCTAGCCATATGTTACTTCGCCTG ACGTTT
glnA	fwd	CAGCAGCGCGAAAACCTGTATTTCAGAGCTCCG CTGAACACGTAC
	rev	CCACCAGTCATGCTAGCCATATGTTAGACGCTGTA GTACAGC
pET28c(+)	fwd	ACAGGTTTCGCCGCTGCTGTGATGATG
	rev	CATATGGCTAGCATGACTGGTG

---

**S.Table 1.1 | Primers to isolate pyrG and glnA from *E. coli* genomic DNA**  
 5' fwd primer contain a 6xHis tag and TEV protease site that is added upstream to the gene. Purified PCR products were then Gibson assembled into pET28c(+) linearized by the primers shown.

name	direction	sequence
+CaM	fwd	CGGCGAAAACCTGTATTTCAGAGCTCCTCCAATC TTACCGAAG
	rev	CGGAGCCGCTACCGCCTTAGATAACAAAGCAGCG A
+MT1a	fwd	CGGCGAAAACCTGTATTTCAGAGCATGGCTGATT CTAATTGTGG
	rev	CGGAGCCGCTACCGCCACAATTACAGTTAACCA CAA
glnA (bB)	fwd	GGCGGTAGCGGCTCCGCTGAACACGTACT
	rev	GCTCTGAAAATACAGGTTTCG

**S.Table 1.2 | Calmodulin (CaM) and plant metallothionein (MT1A) were added to constructed pET28c(+) glnA vector using primers shown.** CaM was isolated from yeast genomic DNA, whereas MT1A was codon-optimized and isolated from the synthesized plasmid. Primers contain the appropriate overhangs to Gibson assemble into the constructed pET28c(+) glnA vector which was linearized with the primers shown.

name	direction	sequence
+3xFlag	fwd	CATGACATCGATTACAAGGATGACGATGACAAGTC CGCTGAACACGTACT
	rev	ATCTTTATAATCACCGTCATGGTCTTGTAGTCGC TCTGAAAATACAGGTTTCG s

**S.Table 1.3 | A 3XFlag tag was added to the constructed pET28c(+) glnA vector.** Primers were used to linearize the pET28c(+) vector with each primer overhang encoding half of the 3XFlag sequence. Linearized product was then blunt-end ligated to re-circularize the plasmid.

name	direction	sequence
glnA	fwd	GCTAGCTCAGCCGAACACGTATTAAC
	rev	GGATCCA ACTGAGTAATACAATTCAAATTCAAC
pyrG	fwd	GCTAGCACAA CGAACTATATTTTGACCc
	rev	GGATCCCTTCGCCTGACGTTCTG

**S.Table 1.4 | Primers used to insert bacterial glnA and pyrG genes into pYD1 yeast display vector.** Primers were used to amplify glnA and pyrG with NheI and BamHI overhangs. The pYD1 vector was cut with NheI and BamHI and ligated with the digested pyrG or glnA product.

# References

- [1] Fenglian Fu and Qi Wang. “Removal of Heavy Metal Ions from Wastewaters: A Review”. In: *Journal of environmental management* 92.3 (2011), pp. 407–418.
- [2] Gökhan Ekrem Üstün, Seval Kutlu Akal Solmaz, and Aşkın Birgül. “Regeneration of Industrial District Wastewater Using a Combination of Fenton Process and Ion Exchange—A Case Study”. In: *Resources, Conservation and Recycling* 52.2 (2007), pp. 425–440.
- [3] T Motsi, NA Rowson, and MJH Simmons. “Adsorption of Heavy Metals from Acid Mine Drainage by Natural Zeolite”. In: *International Journal of Mineral Processing* 92.1 (2009), pp. 42–48.
- [4] Vinod Kumar Gupta et al. “Chemical Treatment Technologies for Waste-Water Recycling—an Overview”. In: *RSC Advances* 2.16 (2012), pp. 6380–6388.
- [5] Adib Amini et al. “Environmental and Economic Sustainability of Ion Exchange Drinking Water Treatment for Organics Removal”. In: *Journal of Cleaner Production* 104 (Oct. 1, 2015), pp. 413–421.
- [6] Anthony M. Wachinski. *Environmental Ion Exchange : Principles and Design, Second Edition*. CRC Press, Oct. 3, 2016.
- [7] Stephan Clemens, Michael G Palmgren, and Ute Krämer. “A Long Way Ahead: Understanding and Engineering Plant Metal Accumulation”. In: *Trends in Plant Science* 7.7 (July 1, 2002), pp. 309–315.

- [8] Christopher Cobbett and Peter Goldsbrough. “Phytochelatins and Metallothioneins: Roles in Heavy Metal Detoxification and Homeostasis”. In: *Annual review of plant biology* 53.1 (2002), pp. 159–182.
- [9] Carolina Sousa, Angel Cebolla, and Víctor de Lorenzo. “Enhanced Metalloadsorption of Bacterial Cells Displaying Poly-His Peptides”. In: *Nature Biotechnology* 14.8 (Aug. 1996), pp. 1017–1020.
- [10] Weon Bae et al. “Enhanced Mercury Biosorption by Bacterial Cells with Surface-Displayed MerR”. In: *Appl. Environ. Microbiol.* 69.6 (June 1, 2003), pp. 3176–3180. pmid: 12788714.
- [11] K. Kuroda and M. Ueda. “Bioadsorption of Cadmium Ion by Cell Surface-Engineered Yeasts Displaying Metallothionein and Hexa-His”. In: *Applied Microbiology and Biotechnology* 63.2 (July 31, 2003), pp. 182–186.
- [12] Mehran Pazirandeh, Bridget M. Wells, and Rebecca L. Ryan. “Development of Bacterium-Based Heavy Metal Biosorbents: Enhanced Uptake of Cadmium and Mercury by Escherichia Coli Expressing a Metal Binding Motif”. In: *Appl. Environ. Microbiol.* 64.10 (Oct. 1, 1998), pp. 4068–4072. pmid: 9758845.
- [13] Kevin J Waldron and Nigel J Robinson. “How Do Bacterial Cells Ensure That Metalloproteins Get the Correct Metal?” In: *Nature Reviews Microbiology* 7.1 (2009), p. 25.
- [14] Lavinia Liliana Ruta et al. “Heavy Metal Accumulation by *Saccharomyces cerevisiae* Cells Armed with Metal Binding Hexapeptides Targeted to the Inner Face of the Plasma Membrane”. In: *Applied Microbiology and Biotechnology* 101.14 (July 1, 2017), pp. 5749–5763.
- [15] P. Stathi et al. “Heavy-Metal Uptake by a High Cation-Exchange-Capacity Montmorillonite: The Role of Permanent Charge Sites”. In: *Global NEST Journal* 12.3 (Apr. 30, 2013), pp. 248–255.
- [16] MA Barakat. “New Trends in Removing Heavy Metals from Industrial Wastewater”. In: *Arabian Journal of Chemistry* 4.4 (2011), pp. 361–377.

- [17] Michele C. Kieke et al. “Selection of Functional T Cell Receptor Mutants from a Yeast Surface-Display Library”. In: *Proceedings of the National Academy of Sciences* 96.10 (May 11, 1999), pp. 5651–5656. pmid: 10318939.
- [18] Nancy Rutherford and Michael Mourez. “Surface Display of Proteins by Gram-Negative Bacterial Autotransporters”. In: *Microbial Cell Factories* 5 (June 20, 2006), p. 22. pmid: 16787545.
- [19] James G. Robertson. “Determination of Subunit Dissociation Constants in Native and Inactivated CTP Synthetase by Sedimentation Equilibrium”. In: *Biochemistry* 34.22 (June 6, 1995), pp. 7533–7541.
- [20] Michael Ingerson-Mahar et al. “The Metabolic Enzyme CTP Synthase Forms Cytoskeletal Filaments”. In: *Nature Cell Biology* 12.8 (Aug. 2010), pp. 739–746.
- [21] Ji-Long Liu. “The Enigmatic Cytoophidium: Compartmentation of CTP Synthase via Filament Formation”. In: *BioEssays* 33.3 (Mar. 1, 2011), pp. 159–164.
- [22] Wendy C. Carcamo et al. “Induction of Cytoplasmic Rods and Rings Structures by Inhibition of the CTP and GTP Synthetic Pathway in Mammalian Cells”. In: *PLOS ONE* 6.12 (Dec. 29, 2011), e29690.
- [23] Ji-Long Liu. “Intracellular Compartmentation of CTP Synthase in Drosophila”. In: *Journal of Genetics and Genomics* 37.5 (May 1, 2010), pp. 281–296.
- [24] Rammohan Narayanaswamy et al. “Widespread Reorganization of Metabolic Enzymes into Reversible Assemblies upon Nutrient Starvation”. In: *Proceedings of the National Academy of Sciences* 106.25 (June 23, 2009), pp. 10147–10152. pmid: 19502427.
- [25] Chalongrat Noree et al. “Identification of Novel Filament-Forming Proteins in *Saccharomyces Cerevisiae* and *Drosophila Melanogaster*”. In: *The Journal of Cell Biology* 190.4 (Aug. 23, 2010), pp. 541–551.

- [26] David Eisenberg et al. “Structure–Function Relationships of Glutamine Synthetases”. In: *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1477.1 (Mar. 7, 2000), pp. 122–145.
- [27] R. C. Valentine, B. M. Shapiro, and E. R. Stadtman. “Regulation of Glutamine Synthetase. XII. Electron Microscopy of the Enzyme from Escherichia Coli”. In: *Biochemistry* 7.6 (June 1968), pp. 2143–2152.
- [28] Richard E. Miller, Emma Shelton, and Earl R. Stadtman. “Zinc-Induced Paracrystalline Aggregation of Glutamine Synthetase”. In: *Archives of Biochemistry and Biophysics* 163.1 (July 1, 1974), pp. 155–171.
- [29] James G. Robertson and Joseph J. Villafranca. “Characterization of Metal Ion Activation and Inhibition of CTP Synthetase”. In: *Biochemistry* 32.14 (Apr. 13, 1993), pp. 3769–3777.
- [30] M. D. Denton and Ann Ginsburg. “Conformational Changes in Glutamine Synthetase from Escherichia Coli. I. Binding of Manganese Ion in Relation to Some Aspects of the Enzyme Structure and Activity”. In: *Biochemistry* 8.4 (Apr. 1969), pp. 1714–1725.
- [31] Steven Knecht et al. “Oligohis-Tags: Mechanisms of Binding to Ni<sup>2+</sup>-NTA Surfaces”. In: *Journal of Molecular Recognition* 22.4 (2009), pp. 270–279.
- [32] John E. Greenleaf, Jin-cheng Lin, and Arup K. Sengupta. “Two Novel Applications of Ion Exchange Fibers: Arsenic Removal and Chemical-Free Softening of Hard Water”. In: *Environmental Progress* 25.4 (2006), pp. 300–311.
- [33] Jacques Haiech et al. “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”. In: *Biochemistry* 20.13 (June 23, 1981), pp. 3890–3897.

- [34] Brian Kelley. “Industrialization of mAb Production Technology The Bioprocessing Industry at a Crossroads”. In: *mAbs* 1.5 (2009), pp. 443–452. pmid: 20065641.