Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supporting Information

Engineering lead-sensing GFP through rational designing

Saravanan Prabhu Nadarajan,[a] Ravikumar Yuvaraj,[b] Kanagavel Deepankumar,[b] Chong-Soon Lee,[b] and Hyungdon Yun*[a]

[a] Department of Bioscience and Biotechnology, Konkuk University, Seoul, Korea.
Email: hyungdon@konkuk.ac.kr
[b] School of Biotechnology, Yeungnam University, Gyeongsan, Gyeongbuk 712-749, Korea.
[*] Corresponding author

1. Materials

PCR reagents, T4 DNA ligase and restriction endonucleases were purchased from Promega (Madison, WI, USA). The host bacterium *Escherichia coli* (*E.coli*) strain XL1-blue (Stratagene, CA, USA) was used for plasmid DNA preparation. Plasmids transformed in *E. coli* were grown aerobically in Luria-Bertani (LB) broth (Difco Laboratories, Detroit, Michigan, USA) or on LB agar plate, supplemented with appropriate antibiotics for the selection of transformants. The pQE-80L plasmid and nickelnitrilotriacetic acid (Ni-NTA) affinity column were purchased from Qiagen (Valencia, CA, USA). All the remaining chemicals isopropyL-D-thiogalactopyranoside (IPTG) were purchased from Sigma-Aldrich, Korea.

2. Construction of plasmids and strains

DNA manipulations were performed according to the procedures described earlier.[1] PCR reaction (50 µl) was carried out with 10 pM of each primer, 50 ng of template DNA, 1X Taq DNA polymerase buffer, 1U of *Taq* DNA polymerase (Promega, Madison, WI, USA), 0.2 mM of each deoxyribonucleotide triphosphates and 1.5 mM $MgCl_2$. DNA amplification was performed in a DNA thermal cycler (Master gradient thermal cycler, Eppendorf, Hamburg, Germany) programmed for an initial denaturation (94ºC for 1 min), followed by 30 cycles of chain reaction for 1 min at 94ºC, 1 min at 60ºC and 0.5 min at 72ºC with an extension at 72ºC for 10 min. GFP was amplified and cloned into pQE-80L. All the constructs were sequenced and confirmed for their target sequences.

3. Molecular modeling studies

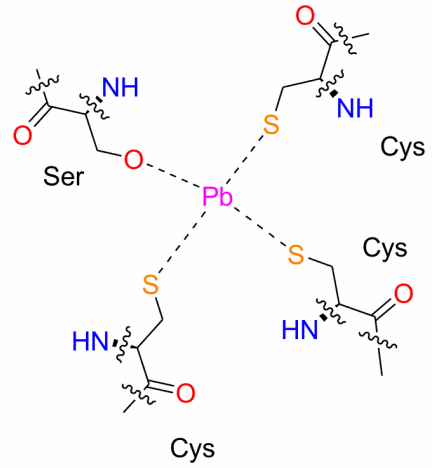
a) Lead binding pattern analysis

Initially, we analyzed the lead binding sites of the earlier reported crystal structures in PDB. We searched for a query ”lead” in the PDB database and obtained the chemical name “lead, Pb^{2+}”. Here, we obtained 40 structures containing lead as free ligand. Among them, 4 structures were DNA and other 36 were protein structures. So we utilized the protein structure and analyzed the lead coordinating sites using PyMol tool.[2] Each protein structures contains one or more lead binding sites. We analysed the interaction with the residues and classfied it into 6 different patterns. Proteins containing lead with no interaction with amino acid residues were ignored.

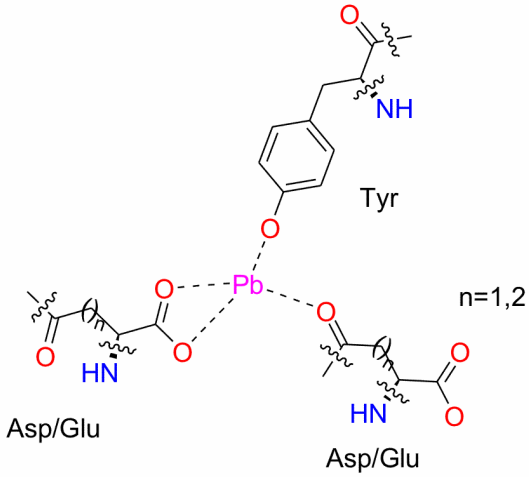
Table S1: Classification of lead binding patterns of protein structure and its lead coordinating residues

Classified based on interaction	PDB ID	Amino acid coordinating lead
Class I	1QNV	Cys133, Cys135, Cys143, Ser179
	1IW7	Cys11, Cys12, Cys58, Cys1112 /Cys58, Cys60, Cys73, Cys76
	1IW7	Cys58, Cys60/Cys60, Cys1204
Class II	1E9N	Tyr171, His309, Asp210, Asn212
	1Na0	Tyr99, Glu98
	2EX3	Glu14, Asp145, Asp169, Tyr148

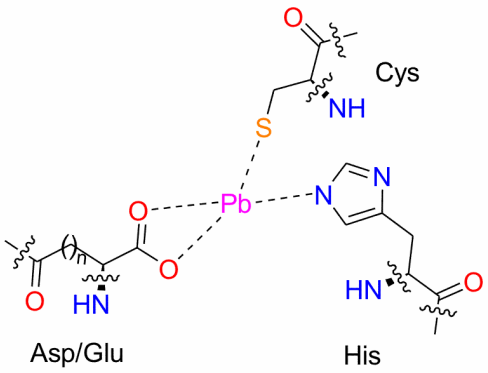
Class III	1QR7	Cys61, His268, Glu302, Asp326
Class IV	1KA4	His269, His273, Glu299
	2FP1	Glu38, Glu41, His151
Class V	1E9N	ASP70, Glu96, Asp308
	1HD7	ASP70, Glu96, Asp308
	1NoY	Glu2, Glu84, Glu87
	1ZHY	Glu204, Arg236 /Asp364
	1ZHZ	Glu204, Arg236 /Asp364
	2G0A	Asp49, Asp238,Asp242
	2V01	Asp22, Asp20, Glu31, Asp24
	3QJK	Asp10, Asp14, Glu21, Ala16
	3QJK	Asp48, Asp46, Glu57, Glu52
	3T8Y	Asp13, Glu60, Asp58
	3TWY	Asp187, Asp246, Asp193, Trp247, Asp248
	1NoY	Glu11, Glu14
	1NA0	Glu61, Glu64
	2CH7	Glu506, Glu247
	2FJ9	Asp69, Glu68
	2FP1	Glu48, Glu51
	4H7X	Asp52, Glu75
	1HQJ	Asp2, Glu10
	2V01	Glu67, Asn60, Thr62
Class VI (single amino acid bound)	1AFV	Glu83
	1FJR	Glu90/ASP139
	1NoY	Glu7
	1SN8	Glu72
	1SYY	Glu270
	1XXA	Glu150
	1ZHW	Asp364
	2ANI	Glu270
	2EX3	Glu221/Asp169
	1V0D	Cys197
	2oQ1	Cys119



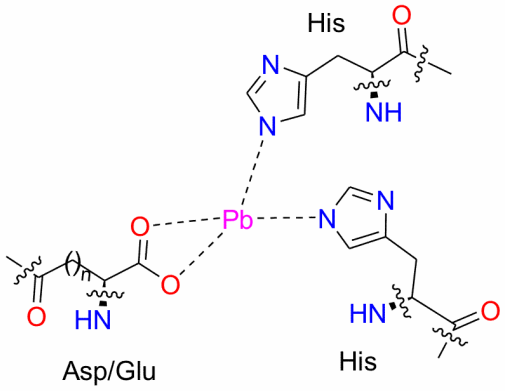
Class I
For eg: PDB ID-1QNV



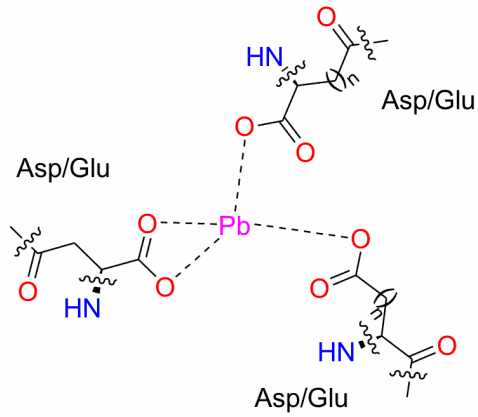
Class II
For eg: PDB ID-1NA0,2EX3



ClassIII
For eg: PDB ID-1QR7

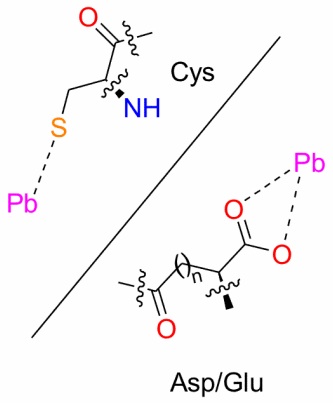


Class IV:
For eg: PDB ID-1KA4



Class V
For eg: PDB ID-1E9N

Pb - Lead



Class VI:
For eg: PDB ID-1AFV,1VoD

Figure S1: Six different lead binding patterns of proteins that co-ordinates lead.

b) GFP chromophore analysis

We utilized the model structure of our GFP variant generated earlier for the copper sensing analysis.[3] Using the model structure of our variant, we analyzed the metal binding sites of the GFP variant. We introduced the *in silico* mutation with the amino acids Cys, His, Asp and Glu at the residues such as Phe145, Ser205, Thr203 using the mutagenesis tool of PyMol. The steric clash was determined and analyzed further.

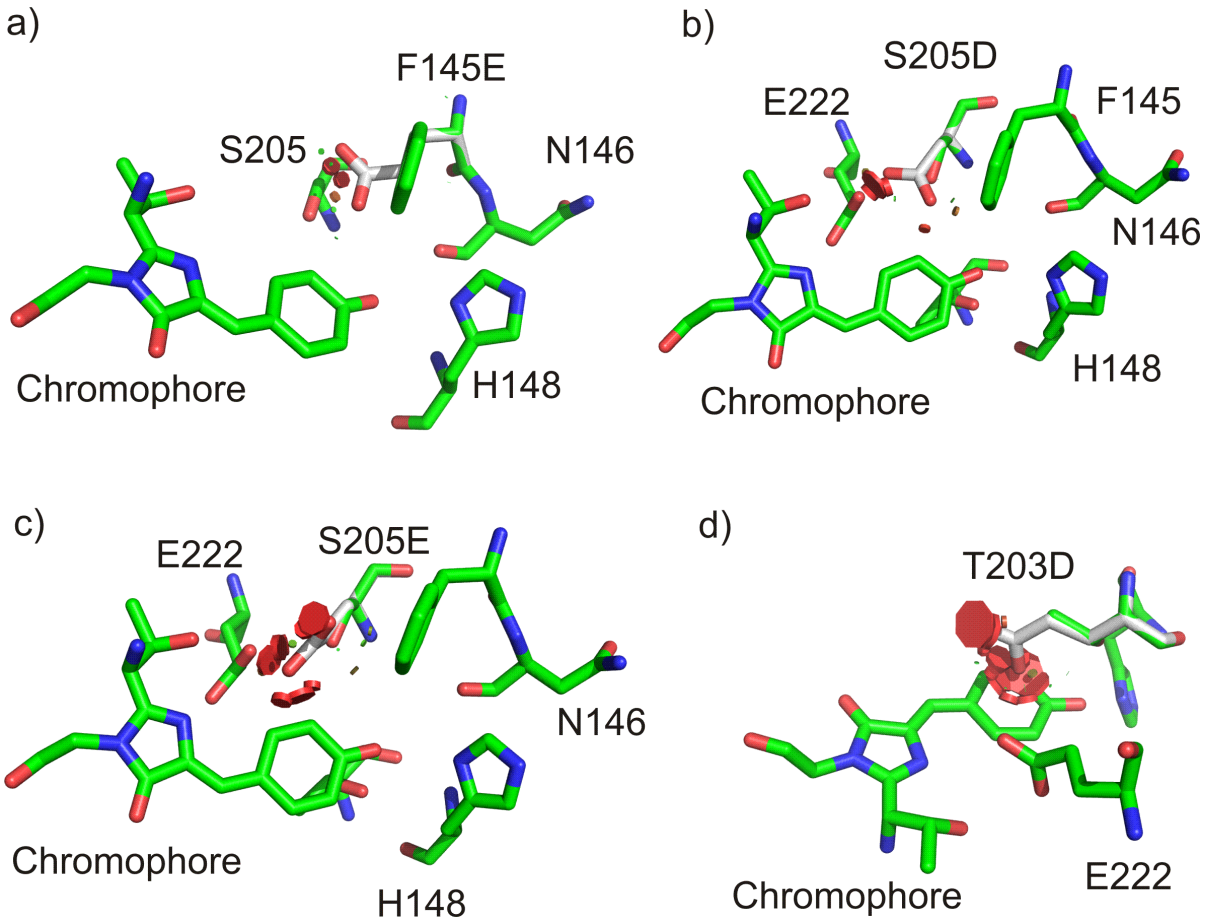


Figure S2: Structure of GFP after *in silico* mutagenesis. Carbon atoms of the mutated residues were represented in gray and other represented as green. a) F145E, Glutamic acid replacing phenylalanine steric clash with Ser205. B,C) In the case of S205D and S205E, the D and E sterically hinder the D222 and Y66. D) Similarly in the case of T203D, T203E, the D and E create a steric clash with Y66.

Table S2: Possible amino acid mutations near the chromophore that enables the protein without deleterious effect and helps lead binding.

Protein	Position	Replacing Amino acid	Possible Fluorescent variant and metal binding activity
GFP	F145	C	Possible mutation
		D	Possible mutation
		E	Steric hindrance With Tyr66,Ser205
		H	Possible mutation
	S205	C	Possible mutation
		D	Steric hindrance With Tyr66, and Glu222
		E	Steric hindrance With Tyr66, and Glu222
		H	Possible but can't help metal binding.
	T203	C	Possible but not help metal binding
		D	Steric hindrance
		E	Steric hindrance
		H	Possible but not help metal binding property
	Y66 (Chromophore tyrosine)	H	Unpredictable

4. Amino acid sequence of GFPHS

MRGSHHHHHHGSMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKITLKL
ICTTGKLPVPWPTLVTT**C**G**Y**GVQCFARYPDHMKRHDFFKSAMPEGYVQERTISFKDDG
TFKTRAEVKFEGDTIVNRIKLKGIDFKEDGNILGHKLEYN**F**NSHKVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVRLPDNHYLS**T**Q**S**VILEDPNEKRDHMVLHEF
VTAAGITHGMDELYK

In this sequence information, 5 amino acids were highlighted in bold red. Cysteine 64, Tyrosine 66, Phenylalanine 145, Threonine 203, and Serine 205.

5. Site directed mutagenesis

Plasmid pQE-80L containing GFP was used for generating different variants such as $GFP_{S205C,Y66H}$ $GFP_{S205C,F145C,}$ $GFP_{S205C,F145D}$, $GFP_{S205C,F145H.}$ The mutant were generated from the GFP by using Quick-change site directed mutagenesis kit (Stratagene, La Jolla, CA) according to the manufacturer's description in the manual. The mutagenesis was confirmed by DNA sequencing analysis at Cosmo Genetech, Daejeon, South Korea.

Table S3: Mutants and its metal sensing property

Protein	C64	Y66	F145	T203	S205	Metal binding
Parent GFP	C	Y	F	T	S	Mercury
$GFP_{Y66H,S205C}$	C	H	F	T	C	Fluorescent loss
$GFP_{F145C,S205C}$	C	Y	C	T	C	Lead and mercury
$GFP_{F145D,S205C}$	C	Y	D	T	C	Copper, Mercury
$GFP_{F145H,S205C}$	C	Y	H	T	C	Mercury
GFP_{C64L}	L	Y	F	T	S	Insensitive to metal
$GFP_{C64L,F145C,S205C}$	L	Y	C	T	C	Lead and mercury
$GFP_{C64L,S205C}$	L	Y	F	T	C	Mercury
$GFP_{C64L,F145C,Cs205A}$	L	Y	C	T	A	Copper and mercury

6. Protein expression and purification

The *E.coli* BL21 (DE3) containing pQE-80L GFP and variants were expressed in LB medium according to previously described protocols.[3] Cells were induced with 1 mM IPTG allowed for 7 h expression. The harvested cells were subjected to centrifugation and stored at -80°C until further use. Briefly, collected cell pellet was suspended in lysis buffer (5 mM immidazole and 50 mM sodium-Phosphate buffer pH 7.4), followed by sonication and centrifuged at 16000 rpm, 4°C for 20 mins. The supernatant was saved as a soluble protein fraction and the pellet was saved as insoluble protein fraction, and analyzed by SDS-PAGE (12% acrylamide gel). The remaining soluble protein fractions were purified by Ni-NTA column chromatography (GE Healthcare Bio-Sciences, Sweden) by standard protocol. Elution fractions were analyzed by SDS-PAGE, and those that were enriched in the desired GFP variants were pooled and dialyzed against 1X phosphate buffered saline. The concentration of protein was quantified using Bradford assay. Fluorescence spectra of GFP variants were recorded on a Perkin Elmer LS-55 fluorescence spectrometer equipped with digital software, Winlab.[3]

7. Metal binding analysis

Metal ion screening was performed for the protein samples prepared in 20 mM MOPS against 0.1 mM metal. Fluorescent measurements were recorded on a Perkin Elmer LS-55 spectrofluorometer using Winlab software.

8. Selectivity study

Stock solutions of heavy metals and metals were prepared of 200 µM and added to GFP variants to a final concentration of 100 µM metal and 1 µM protein. After 1 hour incubation, the fluorescent emission intensity was recorded at 515nm by exciting sample at 430nm using Gemini EM fluorescent micro plate reader (Molecular devices).

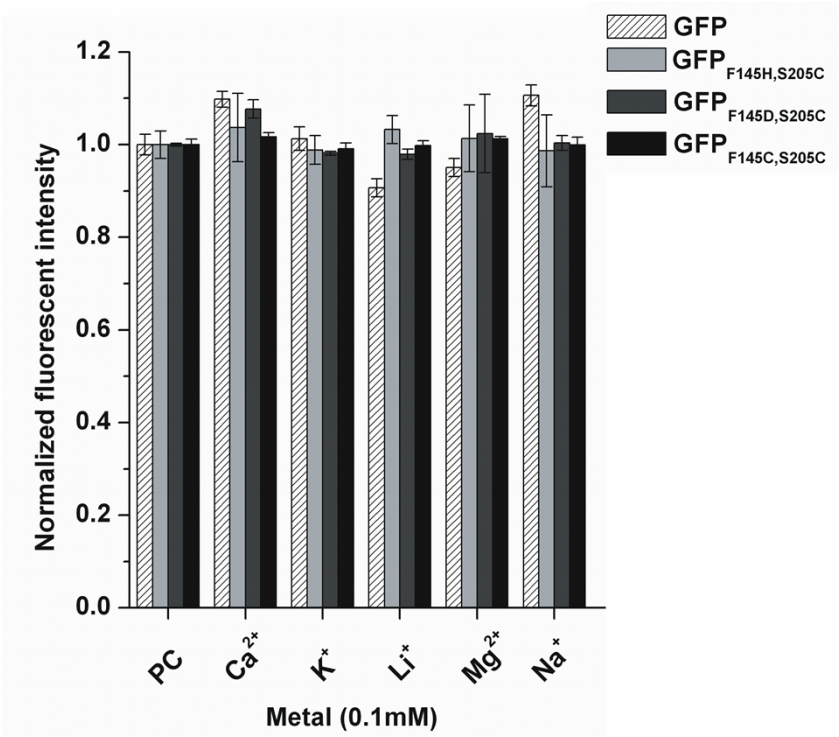


Figure S3. Effect of rest of the metals on GFP variants

9. Lead titration and identification of dissociation constant

A lead titration experiment was carried to determine the apparent binding constant of the protein. To perform the lead binding study, 100 µl of GFP were treated with 100 µl of different concentrations of lead solutions. Fluorescence spectra of these solutions were recorded the resulting emission from 515 excited at 505 nm. This experiment was repeated to calculate standard deviations for each. Fluorescence reading was plotted against Pb^{2+} concentration using the formula mentioned earliar.

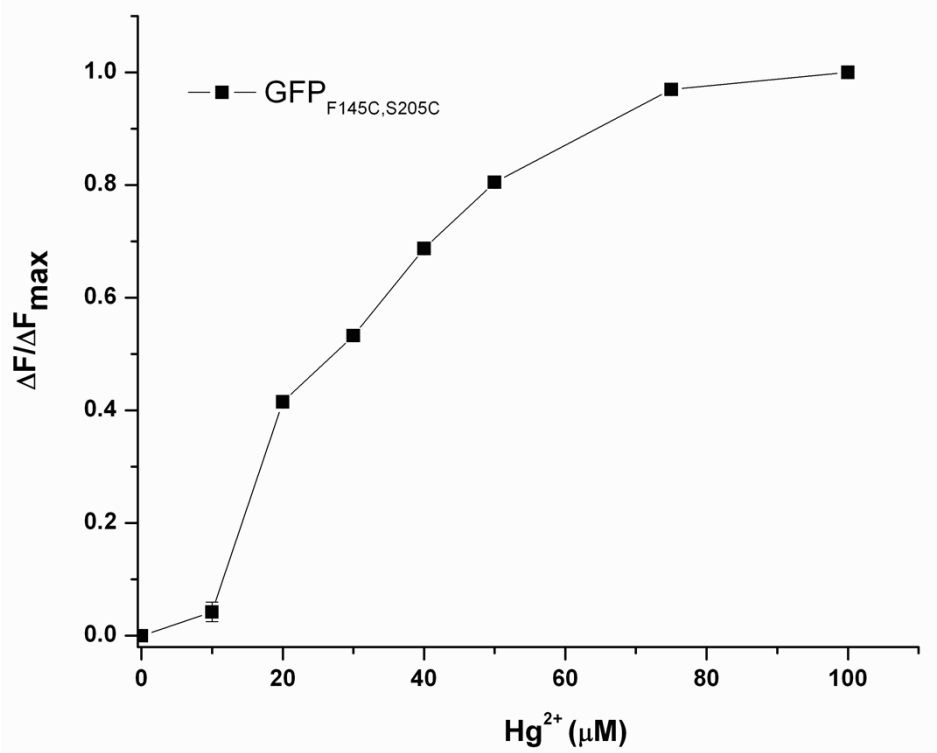


Fig. S4 Fluorescence quenching of 1 μM $GFP_{F145C,S205C}$ in the presence of different concentration of mercury (0-100 μM).

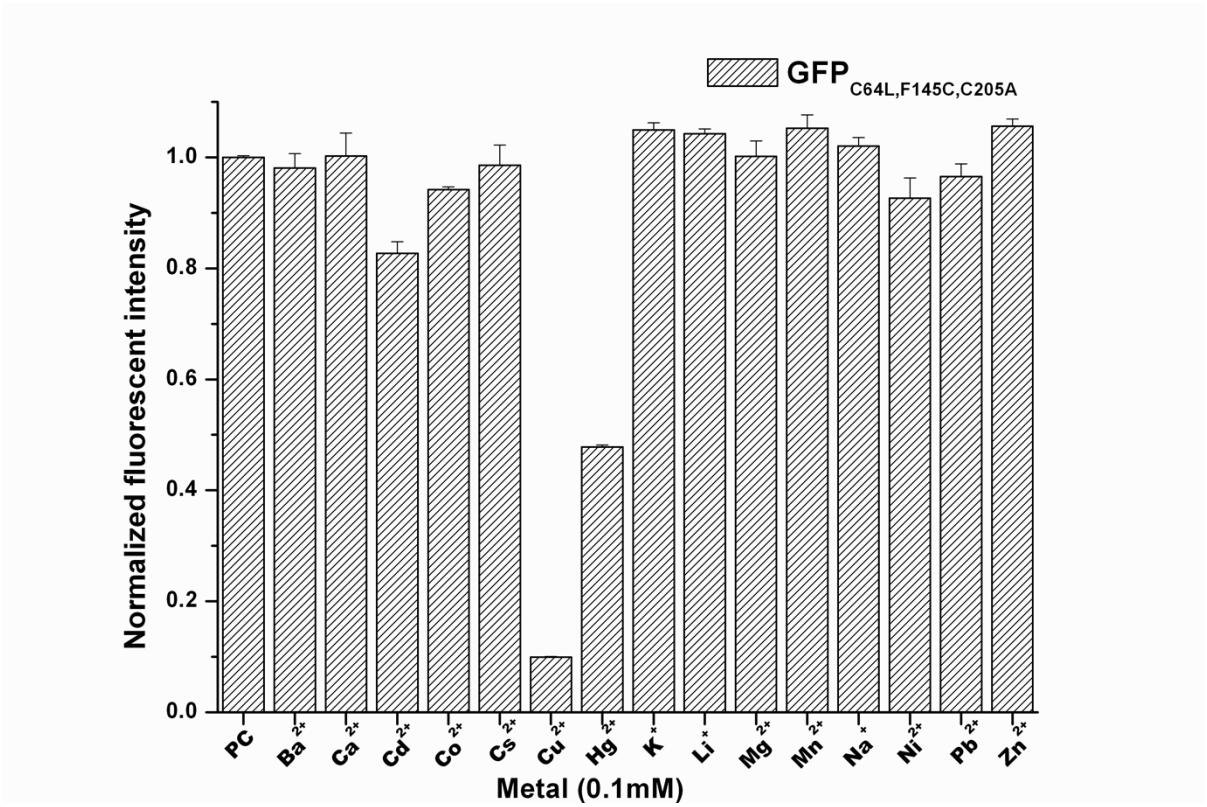


Fig. S5 Metal screening assay for GFP $_{C64L,F145C,\ C205A}$ in various divalent and monvalent metal.

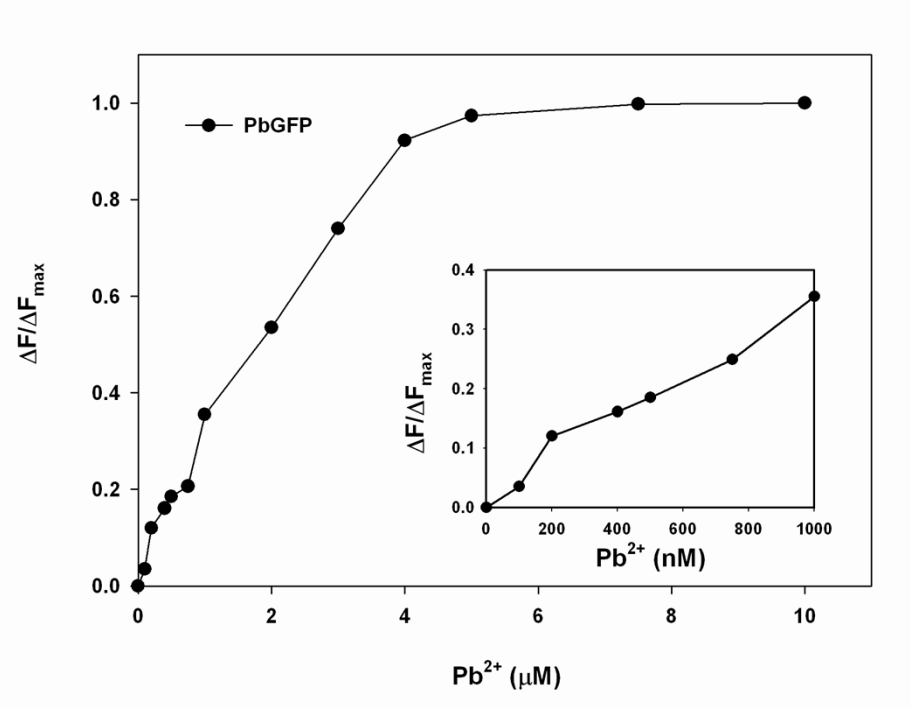


Fig. S6 Calibration curve fo different concentraion of lead in micromolar and nanomolar range against PbGFP .

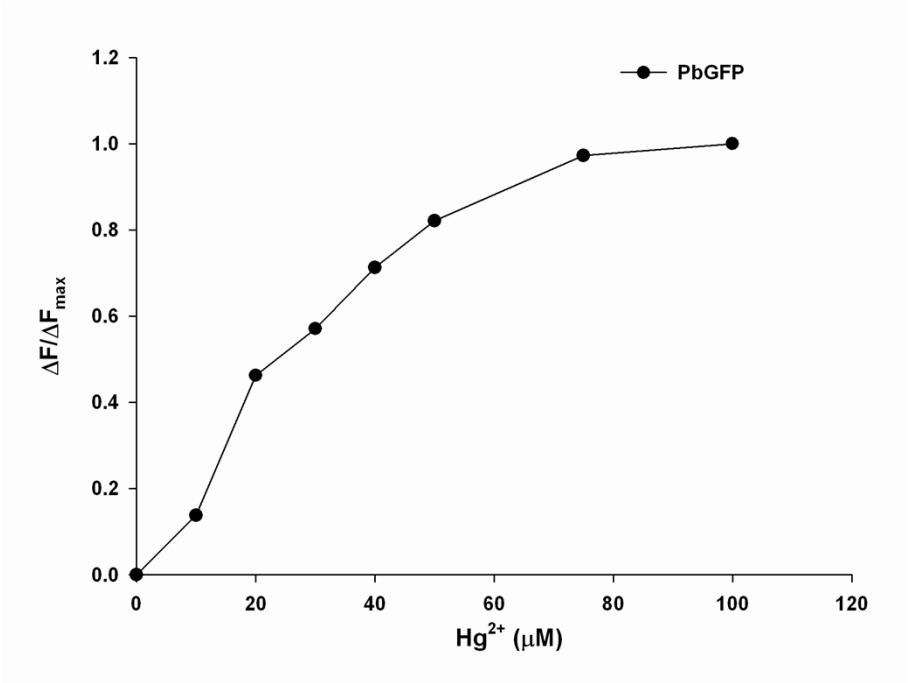


Fig. S7 Effect of different concentration of mercury against PbGFP

10. Effect of EDTA on the metal binding / detecting the reversibility of the sensor

The reversibility of the sensors was assessed by the addition of 1 mM metal chelator, Ethylene diamine tetra acetic acid (EDTA) to the metal treated GFP variants. In addition, we analyzed the effect of EDTA on the lead and mercuryt by simultaneous addition of different concentration of EDTA, 100 μM metal and 1 μM protein

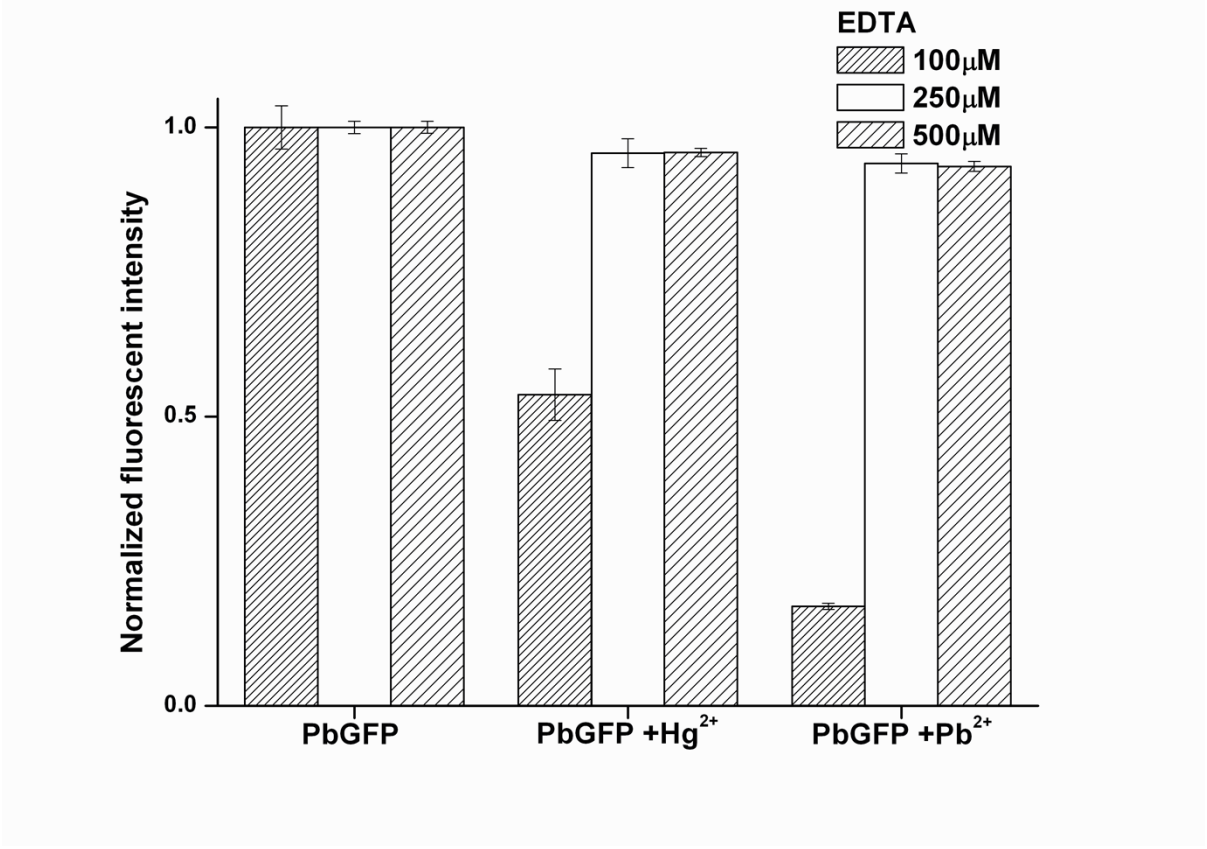


Fig. S8 Effect of different concentration of EDTA on the fluorescence emission intensity of metal treated PbGFP ($GFP_{S205C, F145C,C64L}$)

11. Circular dichroism

A far UV CD spectrum was recorded for the GFP lead sensing variant on a Jasco J-715 spectrophotometer. For the analysis, a volume of 10 μM of protein was placed in a 0.2 cm cell and the CD absorption spectrum was obtained at room temperateure. A volume of 10 μl of 1 mM lead corresponding to final concentraion of 10 μM was added and the absorbance was recorded. 5 Scans were accumulated per spectrum and raw data was processed using Jasco software package. Finally, the obtained results were analysed and graph was drawn using Origin software.

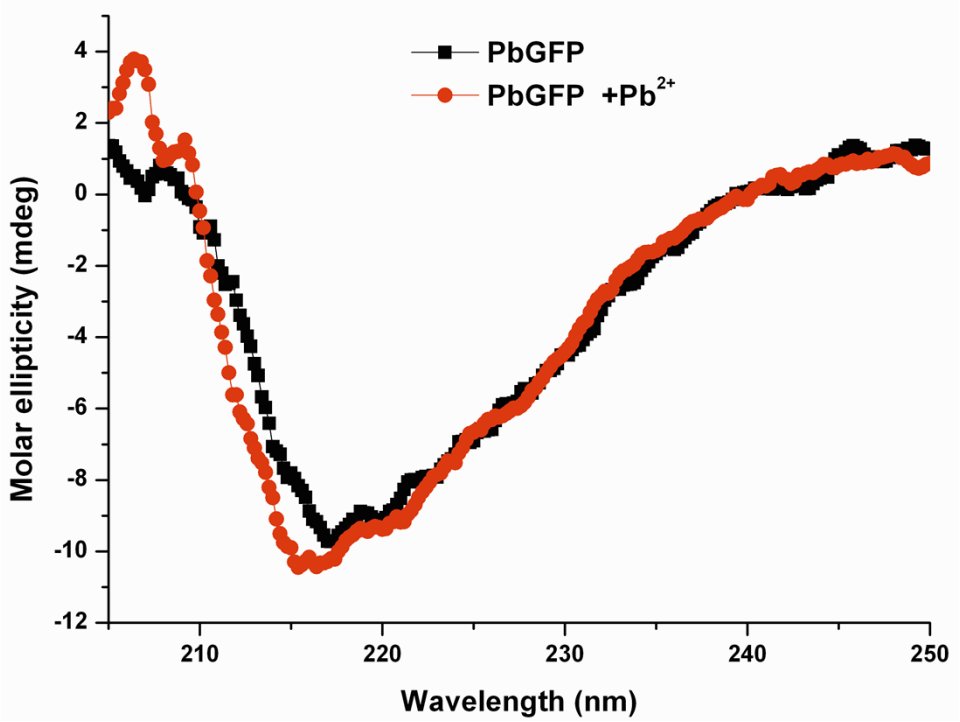


Fig. S9 CD spectral profile of the PbGFP protein in the presence and absence of Pb^{2+}

12. Effect of pH on the metal binding activity

A wide range of pH buffer was prepared and proteins at 1 μM concentration was prepared at respective buffers and treated with lead. Further the fluorescent intensity of the protein prepared at different pH and metal treated variant at different pH was measured and compared.

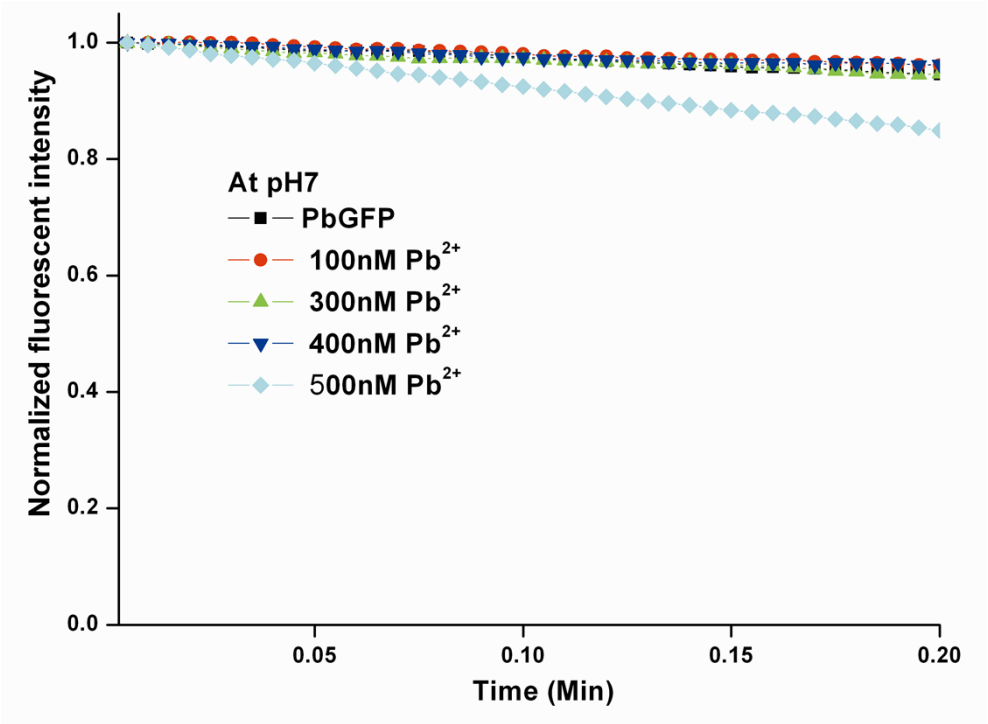


Fig. S10 Time dependent fluorescence quenching of PbGFP at different concentration range from 100 to 500 nM at pH 7.0

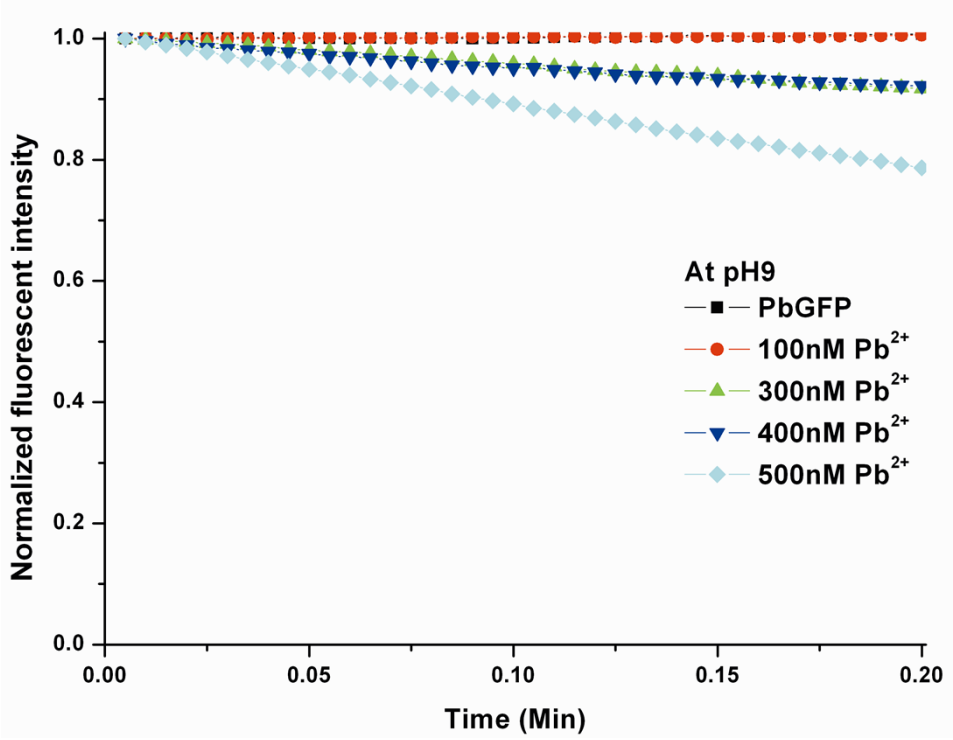


Fig. S11 Time dependent fluorescence quenching of PbGFP at different concentration range from 100 to 500 nM at pH9.0.

13. Bioaccumulation of lead by PbGFP

Metal bioaccumulation assay was performed by culturing the *E.coli* cells in induced and uninduced condition with IPTG. The uninduced *E.coli* cells were utilized as control and the induced *E.coli* cells containg PbGFP were mentioned as PbGFP expressing *E.coli.* Cells were further washed three times with MOPS buffer. Lead nitrate solution was prepared in triple distiled water; 2 OD cells were dissolved in 100 μM lead nitrate solution and then kept in a shaking incubator for 2 hours at 200rpm, 25℃. Further, the cells were centrifuged and the supernatant solution was saved and analysed for the residual Pb^{2+} using atomic absorption spectroscopy. Further the amount of lead accumulated by the cell was calculated in g/DCW.

Reference

[1] J. Sambrook, E. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual. 2nd ed., Cold spring harbor laboratory press, Cold spring harbor, NY, USA, **1989.**

[2] Schrodinger, LLC, **2010.**

[3] N. Ayyadurai, N. Saravanan Prabhu, K. Deepankumar, S. G. Lee, H. H. Jeong, C. S. Lee, H. Yun, Angew Chem Int Ed Engl **2011**, 50, 6534-6537.