Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

A Genetically Encoded Bidentate, Metal-binding Amino Acid

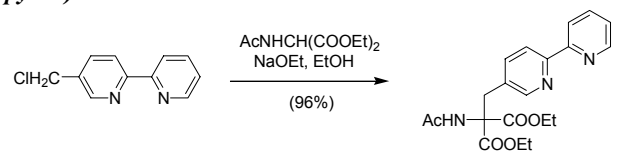
*Jianming Xie, Wenshe Liu, and Peter G. Schultz**

Department of Chemistry and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Correspondence should be addressed to P.G.S. (schultz@scripps.edu).

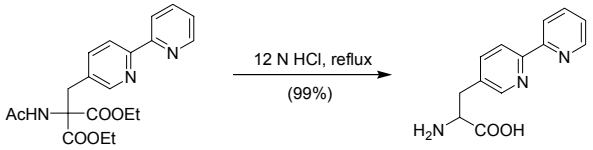
Supplementary methods

Synthesis of bipyridylalanine (BpyAla):



1 2

5-Chloromethyl-2,2'-bipyridine **1** was synthesized from 2-bromopyridine and 2-hydroxy-5-methylpyridine as reported[1]. To a solution of diethyl acetaminomalonate (9.36 g, 43.1 mmol) and NaOEt (2.93 g, 43.1 mmol) in EtOH (80 mL, anhydrous) was added compound **1** (7.06 g, 34.6 mmol). The reaction mixture was heated to reflux overnight. Solvent was evaporated *in vacuo* and the residue was purified by silica gel flash column chromatography (Hexane – EtOAc, 2:1 to 1:1; then CH_2Cl_2 – MeOH, 8:1) to give **2** as a white solid (12.7 g, 96%). **2**: R_f = 0.4 (silica gel, 10:1 CH_2Cl_2:MeOH); ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.76 (d, *J* = 5.0 Hz, 1 H), 8.35 (s, 1 H), 8.32 (dd, *J* = 22.0, 8.0 Hz, 2 H), 7.81 (dt, *J* = 8.5, 2.5 Hz), 7.29 – 7.32 (m, 1 H), 6.60 (s, 1 H), 4.20 – 4.40 (m, 4 H), 3.74 (s, 2 H), 2.07 (s, 3 H), 1.31 (t, *J* = 7.0 Hz, 6 H); ESI-MS: *m/z* 386 (M + H$^+$).



2 3

A suspension of compound **2** (12.6 g, 32.7 mmol) in aqueous HCl (85 mL, 37 % in water) was heated to reflux overnight. The solvent was evaporated *in vacuo* to give **3** [2] as a white HCl salt (11.4 g, 99%), which was used in subsequent biological experiments without further purification.

Construction of libraries of MjTyrRS mutants.

Plasmid pBK-JYRS[3] encodes the wild-type *Mj*TyrRS under the control of *E. coli* GlnRS promoter and terminator, a kanamycin resistant marker and a ColE1 origin of replication. To construct the plasmid DNA library encoding a large number of *Mj*TyrRS variants, alanine mutations were first introduced at the intended randomization sites by overlap extension PCR[4]. Several fragments were first amplified by PCR from pBK-JYRS (*vide infra*). These fragments carried the intended alanine mutation(s) and sequences that overlapped adjacent nucleotides. The products were then purified by electrophoresis on a 2% agarose/TAE gel, diluted, combined, and subjected to overlap extension PCR with two flanking primers JX035NdeI (5'-GAATCCCATATGGACGAATTTG-3') and JX045PstI (5'-TTGAAACTGCAGTTATAATCTCTTTC-3'). The final fragment was digested with the restriction enzymes *Nde*I and *Pst*I, gel-purified, and ligated back into the pBK vector to afford plasmid pBK-JYRS-ALA which encodes the inactive alanine-mutant of *Mj*TyrRS. Using pBK-JYRS-ALA as the template and doped oligonucleotide primers containing NNK codon(s) (*vide infra*), the same procedures described above were repeated to introduce the randomized triplet codon (NNK) at the intended mutation sites. The ligation products (pBK-lib1) that encode library-1 of *Mj*TyrRS variants were then electroporated into *E. coli* DH10B cells (Invitrogen). Electroporated cells were recovered in SOC medium for 60 min at 37°C, and then transferred into a 2 L 2YT medium with kanamycin (50 μg/mL) and were incubated at 37°C to OD_{600nm} ~1.0. To calculate the library size, 1 μL recovered SOC culture was removed and was subjected to serial dilutions in 2YT, and then plated on LB agar plates with kanamycin (50 μg/mL) and grown overnight in a 37°C incubator. Based on the colony numbers on these plates, library-1 contained approximately 1.04×10^9 independent transformants. DNA sequencing of 20 clones did not reveal any significant bias at the randomization site. Plasmid DNA (pBK-lib1) that encodes the *Mj*TyrRSs library was then purified from the amplified 2YT culture by QIAfilter Maxiprep Kit, and used in the subsequent genetic selections whose protocols were described elsewhere [5].

The following primers (all from 5' to 3') were used together with the terminal primers (JX035NdeI and JX045PstI) to introduce alanine or NNK mutations (triplet codons underlined) in the construction of pBK-lib1: JX036Ala: TACCACTTGGTTCAAAACCTATCGCAGCAGATTTTTCATC; JX037NNK: TACCACTTGGTTCAAAACCTATMNNAGCAGATTTTTCATCTTTTTTTAAAACCTCTC; JX038: GGTTTTGAACCAAGTGGTAAAATAC; JX039Ala: CTCCTTTCTGGTTTAAATAGGCCGCTAAATCAGCCAACGCTATAATTATATCAAATCCAGC; JX040NNK:

CTCCTTTCTGGTTTAAATAGGCMNNTAAATCAGCCAAMNNTATAATTATATCAAATCCAGC; JX041: GCCTATTTAAACCAGAAAGGAG; JX042Ala: CCAACTGCAACATCAACGCCCGCATAATGAATCGCATTAACCGCCATTATTGGATAGATAACTTC; JX043NNK: CCAACTGCAACATCAACGCCMNNATAATGAATMNNATTAACMNNCATTATTGGATAGATAACTTC; JX044: GGCGTTGATGTTGCAGTTGG. The terminal primers and primers to introduce NNK mutations were purified by PAGE. The primers to introduce the alanine mutation were used directly without purification. PfuUltra high fidelity DNA polymerase (Stratagene) was used for all PCR reactions according to the manufacturer's protocol. Library-2 and library-3 of *Mj*TyrRS mutants were generated using the same strategy.

Z-domain seqence.
MTSVDNXINKEQQNAFYEILHLPNLNEEQEDAFIQSLKDDPSQSANLLAEAKKLNDAQAPKGSHHHHHH

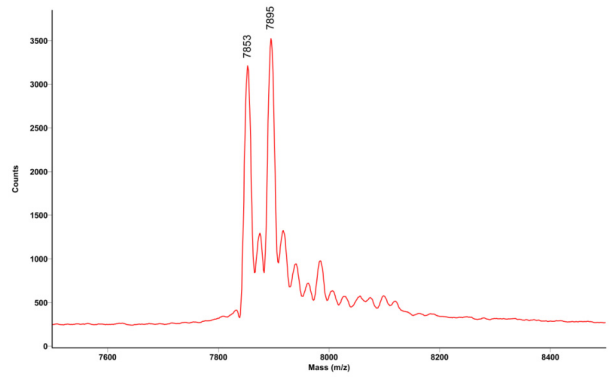
X-ray crystallographic analysis of the BpyAlaRS1/BpyAla complex.

DNA fragment encoding the mutant synthetase BpyAlaRS1 was inserted into the NdeI/XhoI sites of expression vector pET22b to afford plasmid pET22b-BpyAlaRS1. This plasmid was transformed into BL21(DE3) competent cells (Stratagene). The expression and purification of the mutant synthetase followed a similar protocol to that described in previous work[6]. BpyAlaRS1 crystals were grown by hanging drop vapour diffusion method with a 1:1 mixture of ~15mg/ml protein and mother liquor containing 100 mM Tris (pH 7.5-9.0), 5% PEG8K, 16-20% PWG300 and 10% glycerol at 22ºC. X-ray diffraction data of BpyAla were collected at beamline 5.0.3 of the Advance Light Source of Lawrence Berkeley National Laboratory at a wavelength of 1.0 Å to a maximum Bragg spacing to 1.97 Å. The data were reduced and scaled using the HKL2000 package. The structure of BpyAlaRS1 was solved by molecular replacement using wild type *M. jannaschii* tyrosyl-tRNA synthetase (pdb entry as 1J1U) as a search model in which water molecules and other heteroatoms were deleted and processed by Amore of the CCP4 package. Structure refinement was carried out by the combination of Refmac5 of the CCP4 program suite and the CNS program suite. In the refinement process, the program coot in the CCP4 program suite was used for model building, ligand and water finding, and real space refinement of side chains and zones. Biphenylalanine was used as the initial ligand to fit into the difference electron density map at the active site. After all the water molecules were modeled, the ligand was changed to BpyAla in which two pyridine nitrogen atoms were assigned on the basis of their close contact to two water molecules that make strong hydrogen bonding interactions.

References:

[1] S. A. Savage, A. P. Smith, C. L. Fraser, *J. Org. Chem.* **1998,** 63, 10048-10051.
[2] B. Imperiali, T. J. Prins, S. L. Fisher, *J. Org. Chem.* **1993,** *58*, 1613-1616.
[3] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001,** *292*, 498-500.
[4] R. J. Pogulis, A. N. Vallejo, L. R. Pease, *Methods Mol. Biol.* **1996,** *57*, 167-176.
[5] J. Xie, P. G. Schultz, *Methods* **2005,** *36*, 227-238.
[6] J. M. Turner, J. Graziano, G. Spraggon, P. G. Schultz, *Proc. Natl. Acad. Sci. U S A* **2006,** *103*, 6483-6488.
[7] Z. Zhang, L. Wang, A. Brock, P. G. Schultz, *Angew. Chem. Int. Ed. Engl.* **2002,** *41*, 2840-2842.
[8] L. Wang, Z. Zhang, A. Brock, P. G. Schultz, *Proc. Natl. Acad. Sci. U S A* **2003,** *100*, 56-61.
[9] H. Zeng, J. Xie, P. G. Schultz, *Bioorg. Med. Chem. Lett.* **2006,** *16*, 5356-5359.

a



b

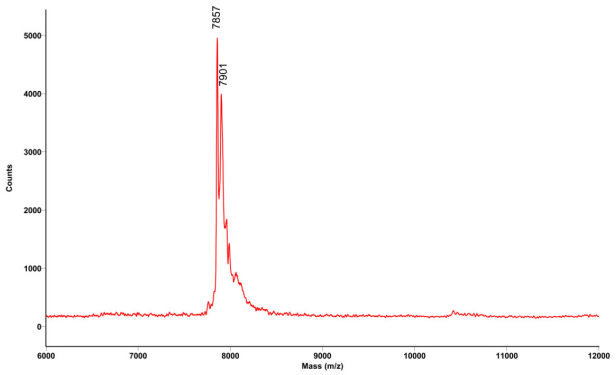


Figure S1. MALDI-TOF-MS analysis of Lys7BipAla and Lys7BpyAla mutants of Z-domain protein. (**a**) The mass spectrum of Lys7BipAla mutant shows two major peaks, corresponding to the mutant Z-domain protein without the first methionine moiety ($M_{Experimental}$ = 7853 Da, $M_{Theoretical}$ = 7853 Da), and the mutant Z-domain protein without the first methionine moiety in acetylated form ($M_{Experimental}$ = 7895 Da, $M_{Theoretical}$ = 7895 Da). (**b**) The mass spectrum of Lys7BpyAla mutant also shows two major peaks, corresponding to the mutant Z-domain protein without the first methionine moiety ($M_{Experimental}$ = 7857 Da, $M_{Theoretical}$ = 7855 Da), and the mutant Z-domain protein without the first methionine moiety in acetylated form ($M_{Experimental}$ = 7901 Da, $M_{Theoretical}$ = 7897 Da). These forms of Z-domain proteins have also been observed in the wild-type and other mutant Z-domain proteins[7-9].

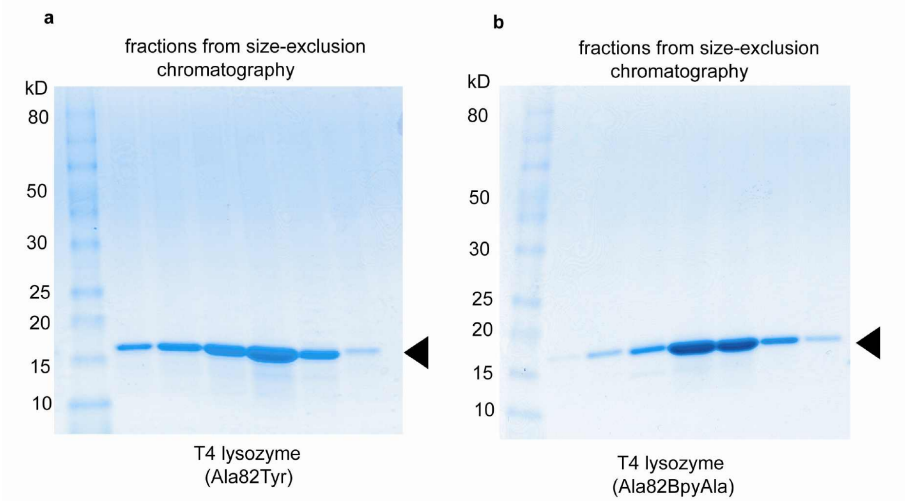


Figure S2. (**a**) Ala82Tyr and (**b**) Ala82BpyAla mutants of T4 lysozyme were purified by gel-exclusion chromatography and the peak fractions were characterized by SDS-PAGE analysis and GelCode staining (Pierce).

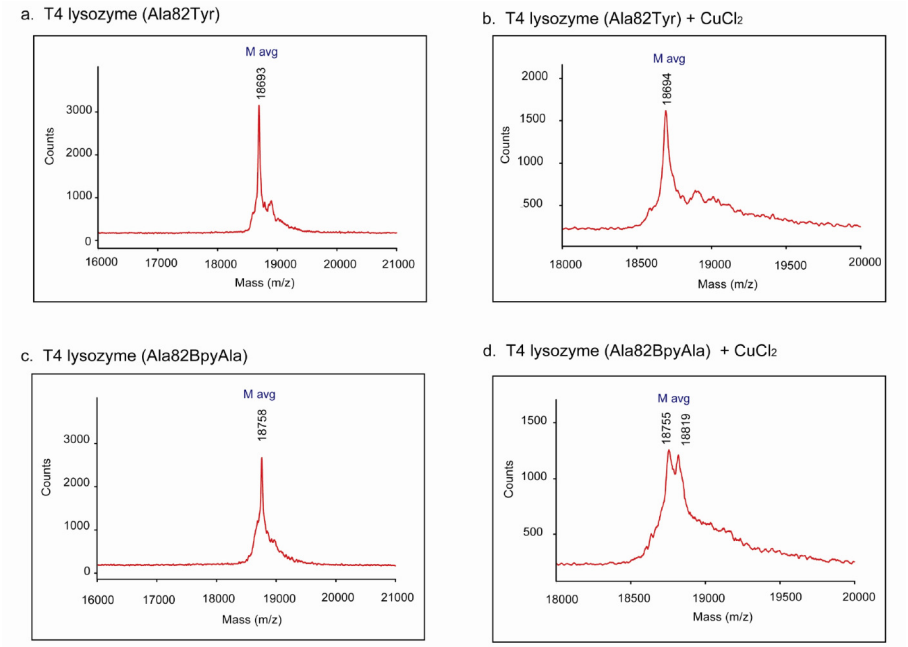


Figure S3. MALDI-TOF-MS analysis of Ala82Tyr and Ala82BpyAla mutants of T4 lysozyme in the presence and absence of $CuCl_2$. (**a**) Ala82Tyr mutant: $M_{Experimental}$ = 18693 Da, $M_{Theoretical}$ = 18694 Da. (**b**) the mixture between Ala82Tyr mutant and $CuCl_2$ (1:3) appeared as one major mass peak at m/z 18694, corresponding to the metal-free molecular ion. (**c**) Ala82BpyAla mutant: $M_{Experimental}$ = 18758 Da, $M_{Theoretical}$ = 18756 Da; (**d**) the mixture between Ala82BpyAla mutant and $CuCl_2$ (1:3) appeared as double peaks at m/z 18755 and 18819, corresponding to the Cu^{2+}-free molecular ion and Cu^{2+}-bound species, respectively.

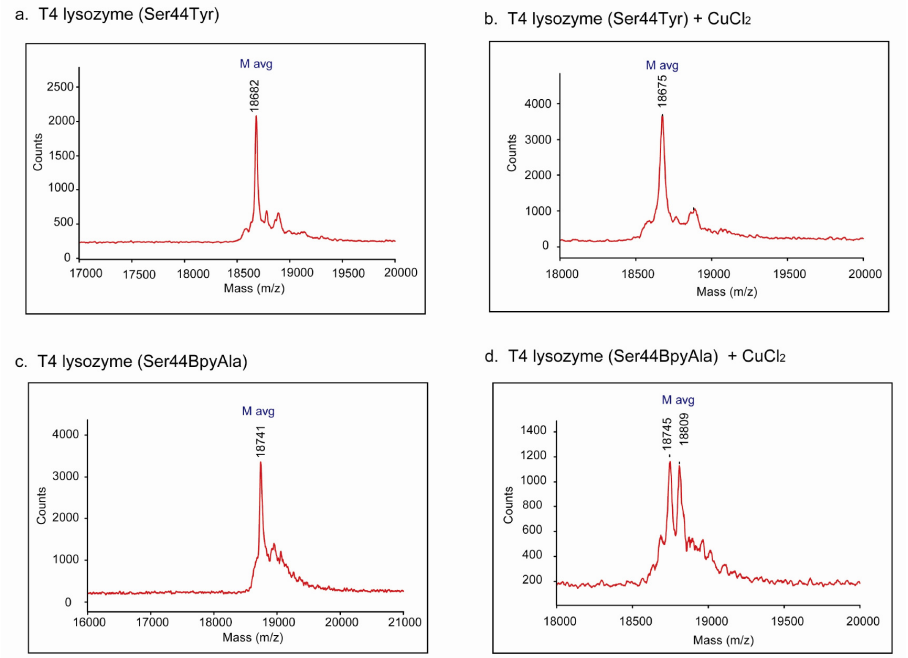


Figure S4. MALDI-TOF-MS analysis of Ser44Tyr and Ser44BpyAla mutants of T4 lysozyme, in the presence or absence of $CuCl_2$. (**a**) Ser44Tyr mutant: $M_{Experimental}$ = 18682 Da, $M_{Theoretical}$ = 18678 Da; (**b**) the mixture between Ser44Tyr mutant and $CuCl_2$ (1:3) appeared as one major mass peak at m/z 18675, corresponding to the metal-free molecular ion. (**c**) Ser44BpyAla mutant: $M_{Experimental}$ = 18741 Da, $M_{Theoretical}$ = 18740 Da; (**d**) the mixture between Ser44BpyAla mutant and $CuCl_2$ (1:3) appeared as double peaks at m/z 18745 and 18809, corresponding to the Cu^{2+}-free molecular ion and Cu^{2+}-bound species, respectively.

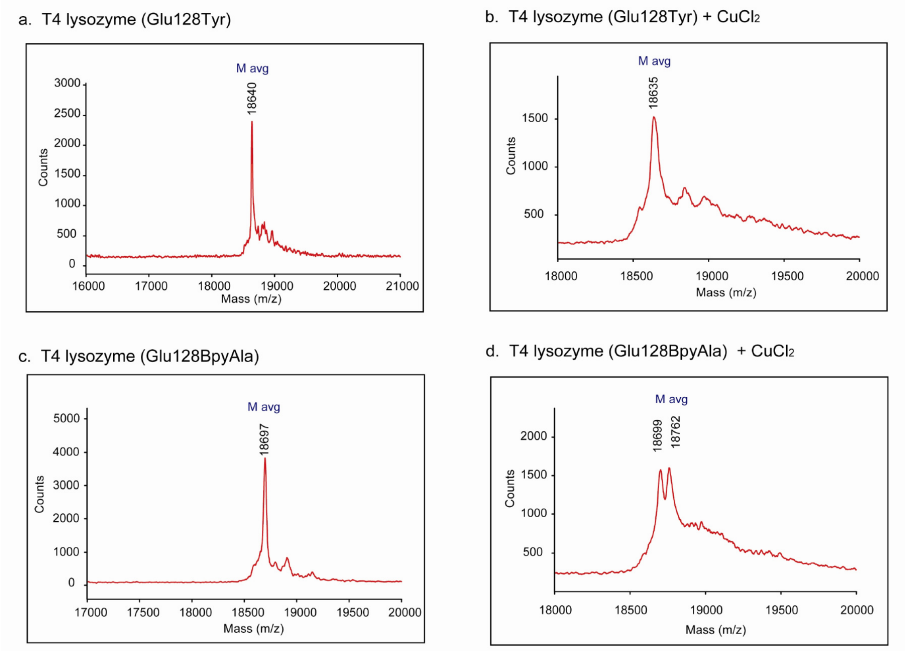


Figure S5. MALDI-TOF-MS analysis of Glu128Tyr and Glu128BpyAla mutants of T4 lysozyme, in the presence or absence of $CuCl_2$. (**a**) Glu128Tyr mutant: $M_{Experimental}$ = 18640 Da, $M_{Theoretical}$ = 18636 Da; (**b**) the mixture between Glu128Tyr mutant and $CuCl_2$ (1:3) appeared as one major mass peak at m/z 18635, corresponding to the metal-free molecular ion; (**c**) Glu128BpyAla mutant: $M_{Experimental}$ = 18697 Da, $M_{Theoretical}$ = 18698 Da; (**d**) the mixture between Glu128BpyAla mutant and $CuCl_2$ (1:3) appeared as double peaks at m/z 18699 and 18762, corresponding to the Cu^{2+}-free molecular ion and Cu^{2+}-bound species, respectively.

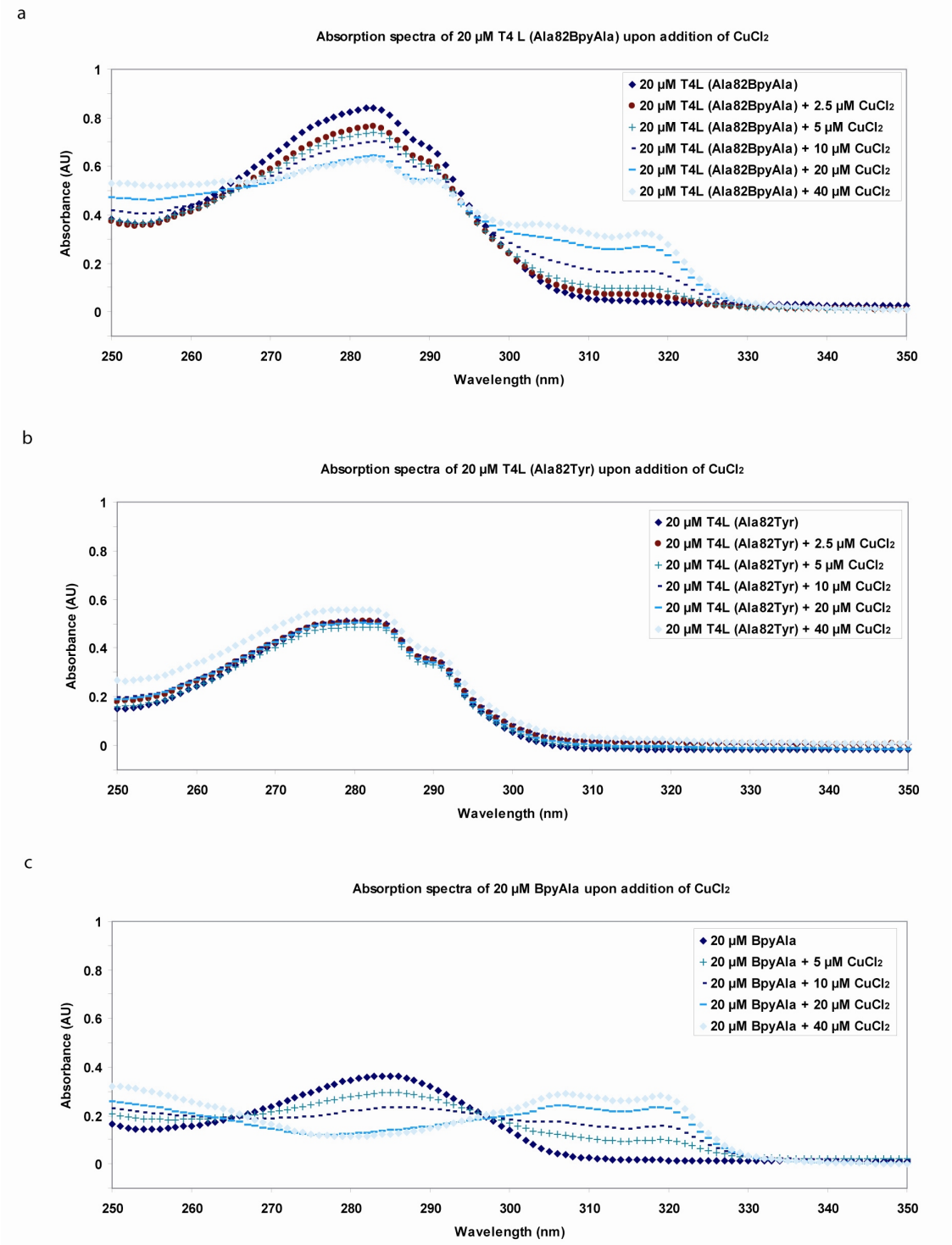


Figure S6. (**a**) The UV-Vis absorption spectra of 20 µM Ala82BpyAla mutant of T4 lysozyme (T4L) upon addition of $CuCl_2$ (from 0 to 40 µM) in 25 mM Tris, 30 mM NaCl (pH = 7.4) showed two new absorption bands at 317 nm and 304 nm, and a decrease in absorption at 283 nm. Such spectral changes are in consistent with the red-shift of $\pi - \pi^*$ transition of the incorporated bipyridyl moiety upon chelation of Cu^{2+} ion; (**b**) In contrast, UV-Vis spectra of 20 µM Ala82Tyr mutant of T4L did not show such changes upon addition of $CuCl_2$; (**c**) For comparison, UV-Vis spectra of 20 µM BpyAla were also measured, which displayed a similar red shift (from 284 nm to 319 nm and 307 nm) as that of Ala82BpyAla mutant of T4L upon addition of $CuCl_2$.