

## **Electronic Supplementary Information**

### **The shielding effect of metal complexes on the binding affinities of ligands to metalloproteins**

Deliang Chen, Yibao Li, Wei Guo, Yongdong Li, Tor Savidge, Xun Li, Xiaolin Fan

## Table of Contents

### Supplementary Texts

1. Examples for calculating water binding energies in Figure 3.....(3)
2. Examples for calculating free energy changes for the formation of metal-ligand complexes in aqueous solution in Figure 4.....(3)

### Supplementary Figures

1. General relationship between the actual binding free energies contributed by metal-ligand interactions and the corresponding binding free energies calculated from the strengths of the metal-ligand interactions. ....(5)
2. Deriving optimal coordination numbers for Zinc complexes in metalloproteins. ....(7)
3. Comparison of binding free energy contributions by the interactions of the carboxyl groups in the binding sites of 3PCF and 3PCG.....(9)
4. Demonstrating why the  $\text{Fe}^{3+}$  in the active site of 3PCG is not able to bind water to form 6-coordinated complex. ....(10)
5. Determining the relative stability of  $\text{Fe}^{3+}$  in the complexes of protocatechuate 3,4-dioxygenase. ....(11)
6. Demonstrating why the  $\text{Fe}^{3+}$  in the active site of 3PCN is not able to bind water to form 6-coordinated complex. ....(12)
7. The formal charges of residue 117 and residue His119 in wild type CAII and in E117Q CAII. ....(13)
8. The zinc-water interaction in wide type CAII is stronger than the zinc-water interaction in E117Q CAII. ....(15)

### Supplementary Tables

9. The actual stepwise water binding energies of metal ions ( $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$  and  $\text{Zn}^{2+}$ ) and the corresponding stepwise water binding energies calculated merely based on M-O interaction strength.....(17)
10. The actual stepwise free energy changes for the formation of metal-ligand complexes in aqueous solution and the corresponding stepwise free energy changes calculated merely based on M-L interaction strength. ....(18)

## Supplementary Texts

**Text S1: The calculation of the  $\Delta E'_n$  values in Figure 3.**  $Zn^{2+}$  as an example to illustrate the calculation of the  $\Delta E'_n$  values shown in Figure 3. As shown in table S1, the actual stepwise water binding energies of  $Zn^{2+}$  obtained from the reference (kcal/mol) are: 101.9 ( $\Delta E_1$ ), 86.6 ( $\Delta E_2$ ), 53.6 ( $\Delta E_3$ ), 41.2 ( $\Delta E_4$ ), 24.0 ( $\Delta E_5$ ), 21.8 ( $\Delta E_5$ ).

$$\text{For } Zn(OH_2): \quad \Delta E_{tot} = \Delta E_1, n = 1; \Delta E'_1 = 101.9$$

$$\text{For } Zn(OH_2)_2: \quad \Delta E_{tot} = \Delta E_1 + \Delta E_2 = 188.5, n = 2; \Delta E'_2 = 94.3$$

$$\text{For } Zn(OH_2)_3: \quad \Delta E_{tot} = \Delta E_1 + \Delta E_2 + \Delta E_3 = 242.1; n = 3; \Delta E'_3 = 80.7$$

$$\text{For } Zn(OH_2)_4: \quad \Delta E_{tot} = \Delta E_1 + \Delta E_2 + \Delta E_3 + \Delta E_4 = 283.3; n = 4; \Delta E'_4 = 70.8$$

$$\text{For } Zn(OH_2)_5: \quad \Delta E_{tot} = \Delta E_1 + \Delta E_2 + \Delta E_3 + \Delta E_4 + \Delta E_5 = 307.3; n = 5; \Delta E'_5 = 61.5$$

$$\text{For } Zn(OH_2)_6: \quad \Delta E_{tot} = \Delta E_1 + \Delta E_2 + \Delta E_3 + \Delta E_4 + \Delta E_5 + \Delta E_6 = 329.1; n = 6; \Delta E'_6 = 54.9$$

**Text S2: The calculation of the  $\Delta G'_n$  values in Figure 4.** The metal complex  $[Ni(NH_3)_n]^{2+}$  is used as an example to illustrate the calculation of  $\Delta G'_n$ . As shown in Table S2, the actual stepwise free energy change for the formation of  $[Ni(NH_3)_n]^{2+}$  ( $\Delta G_n$ , in kcal/mol) are -3.71 ( $\Delta G_1$ ), -2.96 ( $\Delta G_2$ ), -2.27 ( $\Delta G_3$ ), -1.53 ( $\Delta G_4$ ), -0.91 ( $\Delta G_5$ ), and 0.04 ( $\Delta G_6$ ). Below are the calculations of the  $\Delta G'_n$  values of  $[Ni(NH_3)_n]^{2+}$ .

$$\text{For } [Ni(NH_3)]^{2+}: \quad \Delta G_{tot} = \Delta G_1 = -3.71, n = 1, \Delta G'_1 = -3.71;$$

$$\text{For } [Ni(NH_3)_2]^{2+}: \quad \Delta G_{tot} = \Delta G_1 + \Delta G_2 = -6.67, n = 2, \Delta G'_2 = -3.33;$$

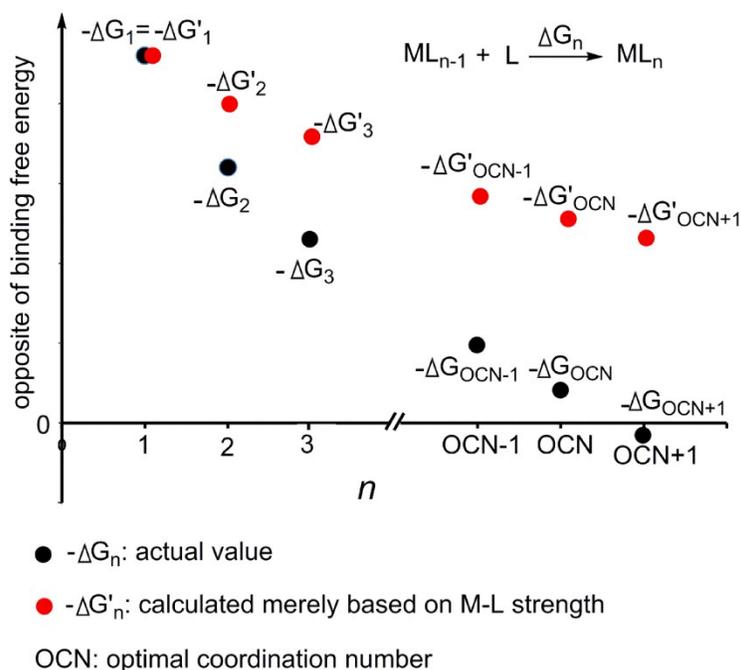
$$\text{For } [Ni(NH_3)_3]^{2+}: \quad \Delta G_{tot} = \Delta G_1 + \Delta G_2 + \Delta G_3 = -8.94, n = 3, \Delta G'_3 = -2.98;$$

$$\text{For } [Ni(NH_3)_4]^{2+}: \quad \Delta G_{tot} = \Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_4 = -10.47, n = 4, \Delta G'_4 = -2.62;$$

For  $[\text{Ni}(\text{NH}_3)_5]^{2+}$ :  $\Delta G_{\text{tot}} = \Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_4 + \Delta G_5 = -11.38$ ,  $n = 5$ ,  $\Delta G'_5 = -2.28$ ;

For  $[\text{Ni}(\text{NH}_3)_6]^{2+}$ :  $\Delta G_{\text{tot}} = \Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_4 + \Delta G_5 + \Delta G_6 = -11.34$ ,  $n = 6$ ,  $\Delta G'_6 = -1.89$ .

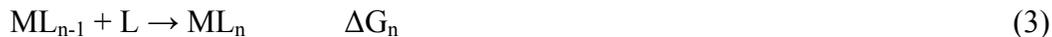
## Supplementary Figures



**Figure S1. General relationship between the actual binding free energies contributed by metal-ligand interactions and the corresponding binding free energies calculated from the strengths of the metal-ligand interactions.** Assume a metal ion M binds with coordinating atom L to form a series metal complexes. The binding processes and the actual stepwise binding free energies are shown in the following equations:



. . . . .



The stepwise binding free energies calculated based on the strengths of the M-L interactions are different from the actual stepwise binding free energies. For example, the two M-L interactions

in  $ML_2$  are identical. The binding free energy calculated from the M-L interaction strength in  $ML_2$ , which is expressed with  $\Delta G'_2$ , is half of the binding free energy of the binding process between M and two L atoms (equation 4). Thus,  $\Delta G'_2 = (\Delta G_1 + \Delta G_2)/2$ .



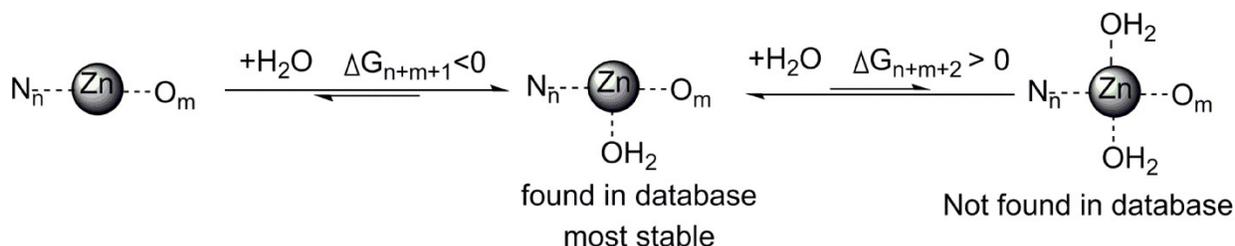
Similarly, the calculated binding free energy for  $ML_{n-1} + L \rightarrow ML_n$ , which is expressed as  $\Delta G'_n$ , is the total binding free energy of  $ML_n$  divided by the number of the number of M-L bonds.

$$\Delta G'_n = (\Delta G_1 + \Delta G_2 + \dots + \Delta G_n)/n \quad (5)$$

Based on equation (5), we get:  $\Delta G'_1$  equals  $\Delta G_1$ ;  $\Delta G'_2$  equals to average value of  $\Delta G_1$  and  $\Delta G_2$ ,  $\Delta G'_3$  equals to average value of  $\Delta G_1$ ,  $\Delta G_2$  and  $\Delta G_3$ , etc.

The actual bonding power (the opposite of the binding free energies) for the metal ion in a metal complex decreases with increasing coordination number (see tables S1 and S2) and becomes zero or negative when the coordination number equals to the optimal coordination number (OCN). The general actual bonding powers for a series of complexes are represented with the black points of this figure. The corresponding bonding powers calculated based on the strengths of M-L interactions, which equals  $-\Delta G'_n$ , are represented with the red points of this figure.

If the bonding power for a metal complex with OCN is slightly less than 0 ( $-\Delta G_{\text{OCN}+1}$  is small negative value), the bonding power calculated based on the interaction strength is about  $-\Delta G_1/2$ . Thus, if the actual bonding power of a metal complex is close to zero, the bonding power calculated merely from the interaction strength is still very strong, indicating the shielding effect is very large and cannot be ignored.



**Figure S2. Deriving optimal coordination numbers for Zinc complexes in**

**metalloproteins.** N and O represent that coordinating atoms of the Zinc complex. N refers

to nitrogen or sulphur atom. O refers to oxygen atom, which may be negatively charged oxygen

atom or neutral oxygen atom. In this figure,  $n$  and  $m$  presents the number of coordinating atoms.

This figure shows that a metal complex  $\text{Zn}(\text{N}_n)(\text{O}_m)$  can bind one water molecule to form

$\text{Zn}(\text{N}_n)(\text{O}_m)(\text{OH}_2)$  but cannot bind two water molecules to form  $\text{Zn}(\text{N}_n)(\text{O}_m)(\text{OH}_2)_2$ . Thus, the

$\text{Zn}^{2+}$  in  $\text{Zn}(\text{N}_n)(\text{O}_m)(\text{OH}_2)$  is sufficiently coordinated and the coordination number (CN) of

$\text{Zn}(\text{N}_n)(\text{O}_m)(\text{OH}_2)$  is the optimal coordination number (OCN) of the zinc complexes containing

$n$  N/S coordinating atoms. The zinc complexes with OCN  $[\text{Zn}(\text{N}_n)(\text{O}_m)(\text{OH}_2)]$  can be found in

PDBbind database by searching the database for such zinc complexes: 1) the zinc complexes

have at least one water molecule to meet the requirement  $\Delta G_{n+m+1} < 0$ ; and 2) the zinc complexes

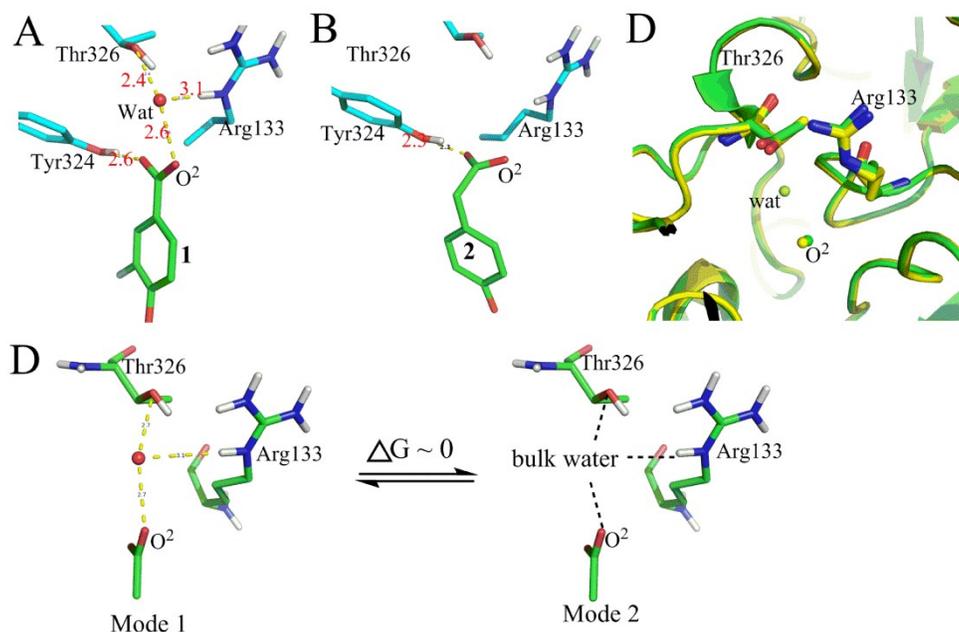
have largest coordination numbers in the zinc complexes containing  $n$  N/S coordinating atoms,

so that  $\Delta G_{n+m+2}$  is larger than 0. By searching such zinc complexes in PDBbind database, the

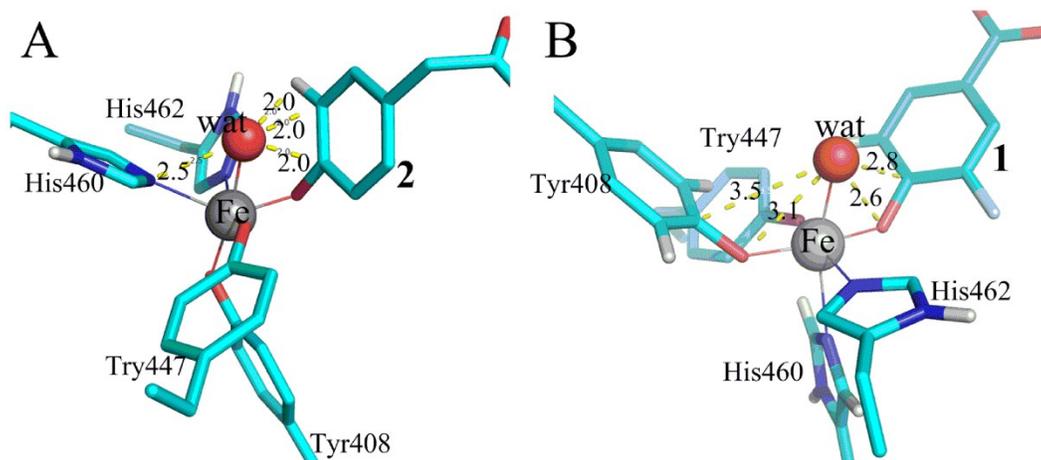
OCNs for various zinc complexes can be derived. The results are shown below:

- 1) 0 N/S: the maximum CN for the zinc complexes containing 0 N/S coordinating atom is 6.
- 2) 1 N/S: the maximum CN for the zinc complexes containing 1 N/S coordinating atom is 6.
- 3) 2 N/S: the maximum CN for the zinc complexes containing 2 N/S coordinating atoms is 6.

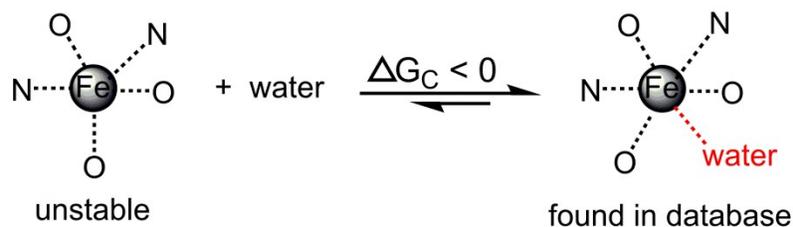
- 4) 3 N/S: if the coordinating atoms also containing negatively charged oxygen atom, the maximum CN is 5; otherwise, the maximum CN is 6.
- 5) 4 N/S: no zinc complexes that contain 4 N/S coordinating atoms also contain water, indicating the binding of such zinc complexes with water is unfavourable and the OCN of the zinc complexes that contain 4 N/S coordinating atoms is  $<5$ . Thus, the OCN of the zinc complexes that contain 4 N/S coordinating atoms is 4.



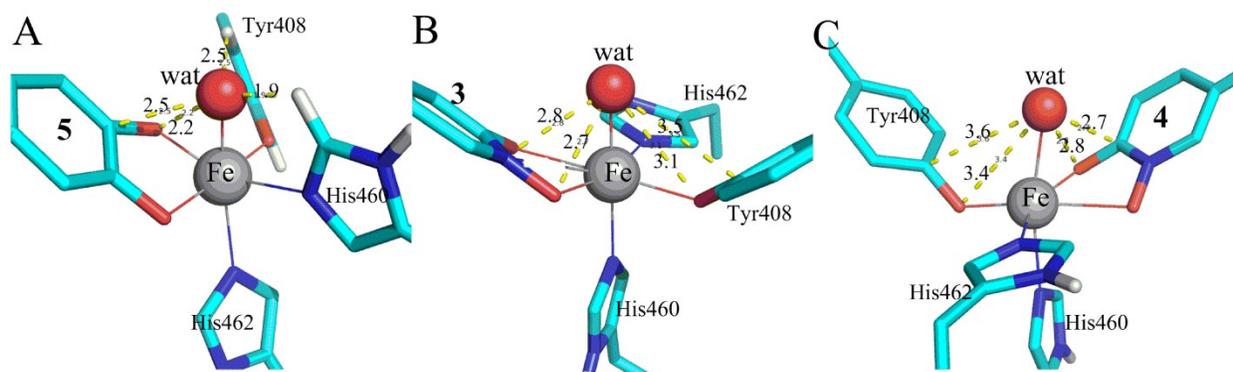
**Figure S3. Comparison of binding free energy contributions by the interactions of the carboxyl groups in the binding sites of 3PCF and 3PCG. (A&B)** The H-bond interactions of the carboxyl groups of the inhibitors. The major difference is that the O<sup>2</sup> atom of **1** interacts with a bridge water molecule, which also interacts with Thr326 and Arg133 via a bridge water (based on pdb: 3PCF), while the O<sup>2</sup> atom of **2** interacts with bulk water because its water accessible surface area (11.3Å<sup>2</sup>) is large enough to interact with bulk water (based on pdb:3PCG). (C) Comparing the environments of the two O<sup>2</sup> atoms (yellow: for inhibitor **1**; green for inhibitor **2**). The environments for the two O<sup>2</sup> atoms are almost identical. (D) Model 1, which is the bridging water interaction model shown in (A), is exchangeable with Model 2, which is the bulk water interaction model shown in (B). As demonstrated in a previous study,<sup>5</sup> the free energy difference between the two models are close to zero. Thus, the binding free energies contributed by the two O<sup>2</sup> atoms of **1** and **2** are almost the same. Thus, the free energy contribution by the carboxyl group of inhibitor **1** is similar to that of inhibitor **2**.



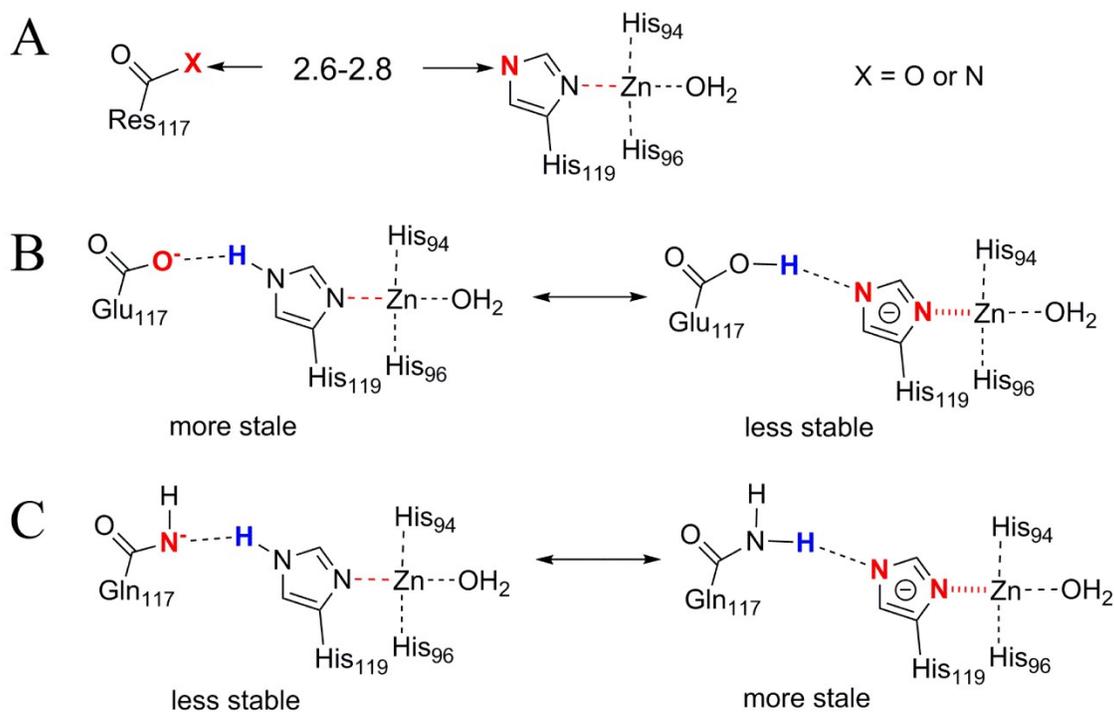
**Figure S4. Demonstrating why the  $\text{Fe}^{3+}$  in the active site of 3PCG is not able to bind water to form 6-coordinated complex.** (A) The active site of 3PCG, which is the crystal structure of protocatechuate 3,4-dioxygenase with inhibitor **2**. In the active site, no metal-bound water is observed. The metal-bound water shown in this Figure is added artificially with the SYBYL program. The added water molecule has very large unfavorable interactions with inhibitor **2** because its distances to two carbon atoms (2.0 Å) of inhibitor **2** are much smaller than the sum of their van der Waals radii (1.52 Å (O) + 1.70 Å (C) = 3.22 Å). (B) The active site of 3PCF, which is the crystal structure of protocatechuate 3,4-dioxygenase with inhibitor **1**. In the active site, a metal-bound water molecule is observed. The shortest distance between the oxygen atom of this water and its non-bound atoms is 2.6 Å, which is only a little smaller than the sum of the van der Waals radii (3.04 Å). This water molecule does not have large unfavorable interactions and is observed in crystal structure. Thus,  $\text{Fe}^{3+}$  in the active site of 3PCG cannot bind water and is 5-coordinated, while the  $\text{Fe}^{3+}$  in the active site of 3PCF can bind water and is 6-coordinated.



**Figure S5. Determining the relative stability of Fe<sup>3+</sup> in the complexes of protocatechuate 3,4-dioxygenase.** The 5-coordinated Fe complexes shown in Figure 7 of the manuscript can bind a water molecule spontaneously because the corresponding 6-coordinated Fe complexes with water in the complexes exist. Thus, the 5-coordinated Fe complexes are hypo-coordinated. However, it does not mean that hypo-coordinated 5-coordinated Fe complexes must bind coordinating atoms to form sufficiently coordinated 6-coordinated Fe complexes because it is possible that there is not enough space for the 5-coordinated Fe complexes to bind water molecules.

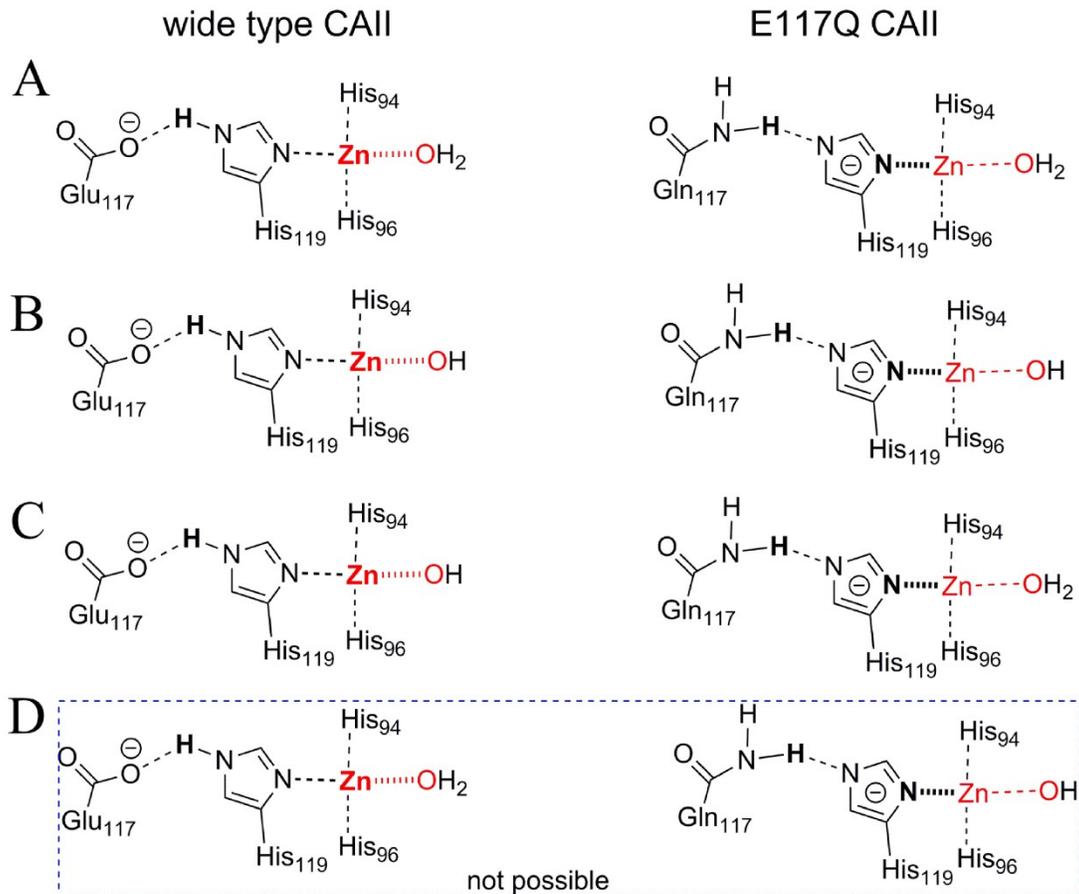


**Figure S6. Demonstrating why the  $\text{Fe}^{3+}$  in the active site of 3PCN is not able to bind water to form 6-coordinated complex.** (A) The active site of 3PCN, which is the crystal structure of protocatechuate 3,4-dioxygenase with inhibitor **5**. In the active site, no metal-bound water is observed. The metal-bound water shown in this Figure is added artificially with the SYBYL program. The added water molecule has very large unfavorable interaction with a non-polar hydrogen atom of His460 because the distance between the atoms ( $1.9 \text{ \AA}$ ) are much smaller than the sum of their van der Waals radii ( $1.52 \text{ \AA} (\text{O}) + 1.20 \text{ \AA} (\text{C}) = 2.72 \text{ \AA}$ ). It also have smaller unfavorable interactions with a carbon of inhibitor **5** because the distance between the two atoms ( $2.5 \text{ \AA}$ ) is obviously smaller than their van der Waals radii ( $1.52 \text{ \AA} (\text{O}) + 1.70 \text{ \AA} (\text{C}) = 3.22 \text{ \AA}$ ). (B&C) The active sites of 3PCJ and 3PCK, which are the crystal structure of protocatechuate 3,4-dioxygenase with inhibitor **3** and **4** respectively. A metal-bound water is observed in each active site. The metal-bound water molecules do not have large unfavorable interactions with their environment. Thus,  $\text{Fe}^{3+}$  in the active site of 3PCN cannot bind water and is 5-coordinated, while the  $\text{Fe}^{3+}$  ions in the active sites of 3PCJ and 3PCK can bind water and is 6-coordinated.



**Figure S7. The formal charges of residue 117 and residue His119 in wild type CAII and in E117Q CAII.** (A) The distance between the oxygen atom (or the nitrogen atom) of residue 117 and the nitrogen atom of His119 in wild type CAII or in E117Q CAII. By analysing the crystal structures of wild type CAII, E117Q CAII, and their complexes with inhibitors, we found that distances range from 2.6 to 2.8Å. It is impossible that both X and His119 are negatively charged because of the large unfavourable negative-negative repulsive interaction. It is also impossible the both X and the nitrogen atom of His119 contain hydrogen atoms because the distance between X and the nitrogen atom is too short to have a hydrogen atom on each atom. Thus, there is exactly one hydrogen atom between X and the nitrogen atom of His119. Either residue 117 or His119 must be negatively charged. (B) For wild type CAII, the negatively charged carboxyl group is more stable than the negatively charged His119. Thus, His119 is

neutral in wild-type CAII. (C) For E117Q CAII, the negative charged His119 is more stable than the negatively amide group because the pKa of the imidazole group in histidine is much lower than that of amide group. Thus, His119 is negatively charged His119 in E117Q CAII.



**Figure 8. The zinc-water interaction in wide type CAII is stronger than the zinc-water interaction in E117Q CAII.** As demonstrated in Figure S7, the His119 in wide type CAII is neutral and the His119 in E117Q CAII is negatively charged. Thus, the zinc in wide type CAII is more positively charged than the zinc ion in E117Q CAII. A zinc bound water molecule could be in its neutral state or in deprotonated state. If the water molecules in both wide type CAII and E117Q CAII are in their neutral states as shown in (A) or in their deprotonated states as shown in (B), the zinc-water interaction in wide type CAII is stronger than the zinc-water interaction in E117Q CAII. Because the zinc in wide type CAII is more positively charged than the zinc ion in E117Q CAII, it is also possible that the water molecule in wide type CAII is in the deprotonated state, while the water molecule in E117Q CAII in the neutral state. In this case,

the zinc-water interaction in wide type CAII is much stronger than the zinc-water interaction in E117Q CAII. It is impossible that the water molecule in wide type CAII is in the neutral state, while the water molecule in E117Q CAII is in the deprotonated state as shown in (D). Thus, no matter the zinc-bound water molecules are in their neutral states and deprotonated states, the zinc-water interaction in wide type CAII is always stronger than the zinc-water interaction in E117Q CAII.

## Supplementary Tables

**Table S1:** The actual stepwise water binding energies of metal ions ( $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$  and  $\text{Zn}^{2+}$ ) [ $\Delta E_n$  for  $\text{M}(\text{OH}_2)_{n-1} + \text{H}_2\text{O} \rightarrow \text{M}(\text{OH}_2)_n$ ] and the corresponding stepwise water binding energies calculated merely based on M-O interaction strength( $\Delta E'_n$ ).\*

n	$\text{Mg}^{2+}$		$\text{Ca}^{2+}$		$\text{Zn}^{2+}$	
	$\Delta E_n$	$\Delta E'_n$	$\Delta E_n$	$\Delta E'_n$	$\Delta E_n$	$\Delta E'_n$
1	81.5	81.5	56.9	56.9	101.9	101.9
2	70.9	76.2	47.5	52.2	86.6	94.3
3	55.1	69.2	42.0	48.8	53.6	80.7
4	43.9	62.9	35.6	45.5	41.2	70.8
5	28	55.9	27.7	41.9	24.0	61.5
6	24.5	50.7	24.7	39.1	21.8	54.9
7	3.9	44.0	13.8	35.5		
8			8.8	32.1		

\*The actual stepwise water binding energies ( $\Delta E$ ) are obtained from reference.<sup>6</sup> The method for calculating the corresponding stepwise water binding energies calculated merely based on M-O interaction strength ( $\Delta E'$ ) is shown in Methods.

**Table S2:** The actual stepwise free energy changes for the formation of metal-ligand complexes in aqueous solution [ $\Delta G_n$  for  $ML_{n-1} + L \rightarrow ML_n$ ]\* and the corresponding stepwise free energy changes ( $\Delta G'_n$ ) calculated merely based on M-L interaction strength.\*\*

n	Ni <sup>2+</sup> /MH <sub>3</sub>		Co <sup>2+</sup> /MH <sub>3</sub>		Cd <sup>2+</sup> /CH <sub>3</sub> COO <sup>-</sup>	
	$-\Delta G_n$	$-\Delta G'_n$	$-\Delta G_n$	$-\Delta G'_n$	$-\Delta G_n$	$-\Delta G'_n$
1	3.71	3.71	2.72	2.72	2.58	2.58
2	2.96	3.34	2.06	2.39	1.21	1.90
3	2.27	2.98	1.27	2.02	0.03	1.27
4	1.53	2.62	0.87	1.73	-0.68	0.78
5	0.91	2.28	0.08	1.40		
6	-0.04	1.89	-1.00	1.00		

\* Metal-bound water molecules are displaced when the M-L bonds are formed.

\*\*The actual stepwise actual stepwise free energy changes ( $\Delta G_n$ ) are calculated from stepwise association constants, which are obtained from references.<sup>7</sup> The corresponding stepwise free energy changes calculated merely based on M-L interaction strength ( $\Delta G'_n$ ) are calculated from actual stepwise actual stepwise free energy changes (see Methods).

1. Wang, R.; Fang, X.; Lu, Y.; Yang, C. Y.; Wang, S. The PDBbind database: methodologies and updates. *J. Med. Chem.* **2005**, *48*, 4111-9.
2. Shrake, A.; Rupley, J. A. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* **1973**, *79*, 351-71.
3. Huang, C. C.; Lesburg, C. A.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. Reversal of the hydrogen bond to zinc ligand histidine-119 dramatically diminishes catalysis and enhances metal equilibration kinetics in carbonic anhydrase II. *Biochemistry* **1996**, *35*, 3439-46.
4. Wang, R.; Lu, Y.; Fang, X.; Wang, S. An extensive test of 14 scoring functions using the PDBbind refined set of 800 protein-ligand complexes. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2114-25.
5. Chen, D.; Li, Y. B.; Zhao, M.; Tan, W.; Li, X.; Savidge, T.; Guo, W.; Fan, X. L. Effective lead optimization targeting displacement of bridging receptor-ligand water molecule. *Phys. Chem. Chem. Phys.* **2018**, DOI: 10.1039/C8CP04118K.
6. Pavlov, M.; Siegbahn, P. E. M.; Sandstrom, M. Hydration of Beryllium, Magnesium, Calcium, and Zinc Ions Using Density Functional Theory. *J. Phys. Chem. A* **1998**, *102*, 219-28.
7. Martell, A. E.; Smith, R. M. *Critical stability constants*. Plenum Press: New York., 1976; p v. <1-6>.