Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information

Effective lead optimization targeted for displacing bridging water molecule

Deliang Chen,* Yibao Li, Mingming Zhao, Wen Tan, Xun Li, Tor Savidge, Wei Guo, Xiaolin Fan*

Table of content

Methods

Supplementary Text

Supplementary Figures

References

I. Methods

Calculation of H-bonding capability: The H-bonding capabilities of hydrogen (H)-bond acceptors and molecules were calculated according to the methods reported in a previous study.[1] The H-bonding capability of H-bond donors (H_{HBD}) in a molecule was calculated from the following equation:

$$H_{HBD} = 0.822\ \mathrm{H_M} - 1.188\ \mathrm{H^{oct}_M} + 1.251 \tag{M1}$$

where H_M, H-bonding capability of the molecule, was calculated from the hexadecane/water partition coefficients as described previously.[1] H^{oct}_M, which was calculated from the 1-octanol/water partition coefficients (P_{oct}) using the same method as for calculating H_M.

Comparison of the water-binding sites at protein-ligand interfaces of 3PCK and 3PCN. The coordinates for the protocatechuate 3,4-dioxygenase complexed with 6-hydroxynicotinic acid N-oxide (3PCK)[2] and protocatechuate 3,4-dioxygenase complexed with (3PCN) were retrieved from the Protein Data Bank (PDB).[3] The "Align Structures by Homology" program implemented in SYBYL (Tripos Inc., St Louis, MO) was used to overlap the water-binding sites at the protein-ligand interfaces of 3PCK and 3PCN. Only the A and M chains of 3PCK and 3PCN were used for alignment, while other chains were deleted because they are far from the water-binding sites for comparison. The structure 3PCK was selected as fixed reference structure, while 3PCN was selected as movable structure. All atoms were used for the least-squares fits. The calculated RMSD after alignment is 0.31. The overlap of the water binding sites is shown in Figure 2C.

Predicting the location of the bridging water molecule in the water binding site of 3PCG. The coordinates for the protocatechuate 3,4-dioxygenase complexed with 3-hydroxyphenylacetate (3PCE),[4] with 3-fluro-4-hydroxybenzoate (3PCF),[4] and 4- hydroxyphenylacetate (3PCG),[4] 2-hydroxyisonicotinic acid N-oxide (3PCJ),[2] and with 6-hydroxynicotinic acid N-oxide (3PCK)[2] were retrieved from the Protein Data Bank (PDB). No bridging water molecule in water binding site of 3PCG is observed. However, bridging water molecules in 3PCE, 3PCF, 3PCJ and 3PCK are observed. Thus, the structure of 3PCG was superimposed with the structures of 3PCE, 3PCF, 3PCJ, and 3PCK by fitting the polar atoms in the water binding sites with the "Fit Atoms" program implemented in SYBYL. The fitting of 3PCG to 3PCF has lowest RMSD value, which suggests that the bridging water molecule in the water binding site of 3PCG is close to the bridging water molecule of 3PCF. Thus, the location of the bridging water molecule of 3PCF was used as the location of the bridging water molecule of 3PCG. The result is shown in Figure S8B.

II. Supplementary Text

Text S1. Free energy change of water freezing at 298K calculated from the thermodynamic data of water: The water freezing free energy change (ΔG) at temperature T can be obtained from the enthalpy change (ΔH) and entropy change (ΔS) of the water freezing process. The ΔH term is only weakly dependent on temperature, but the TΔS term is strongly dependent on temperature because the presence of the T in the term. Assume the water freezing ΔGs at temperature T_1 and T_2 are ΔG_{T1} and ΔG_{T2} respectively. The relationship between ΔG_{T1} and ΔG_{T2} is

$$\Delta G_{T2}/T_2 = \Delta G_{T1}/T_1 + \Delta H(1/T_2 - 1/T_1) \qquad \text{(S1)}$$

Assume T_1 is the water freezing point, which is 273K. Then ΔG_{T1} is zero. Since the enthalpy of fusion for water is 6.007 kJ/mol, thus, ΔH equals -6.007 kJ/mol. The water freezing ΔG at 298K can be obtained from the following equation:

$$\Delta G_{298}/298 = 0/273 - 6.007\ (1/298 - 1/273) \qquad \text{(S2)}$$

From equation (S2), we obtain the water freezing free energy change at 298 K is 0.55 kJ/mol.

III. Supplementary Figures

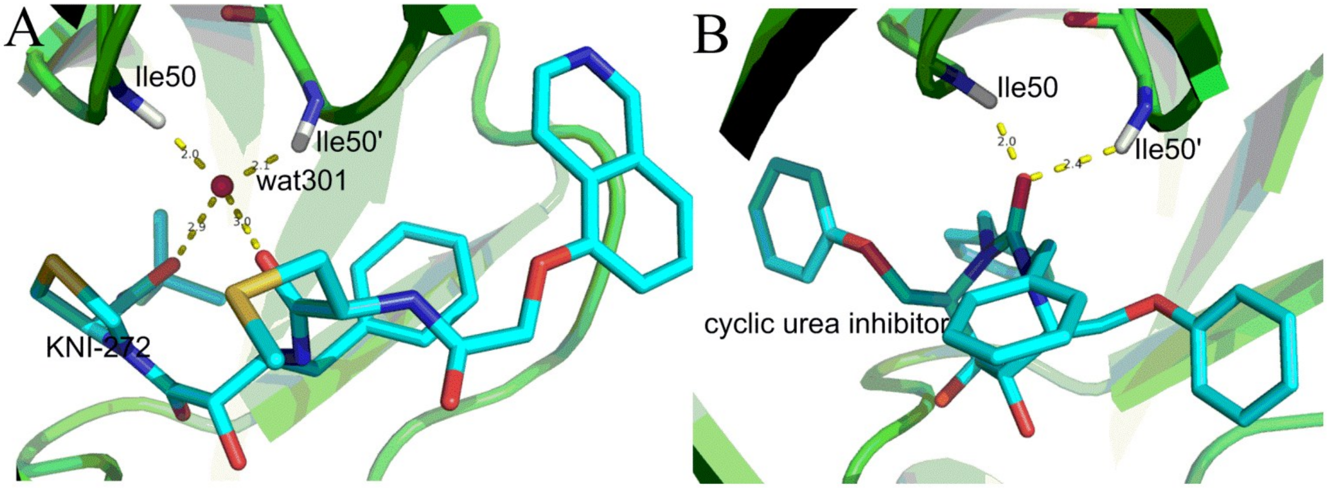

Figure S1. Change in the number of ordered H-bonds resulted from the displacement of the bridging water molecule at the interface of HIV-1 protease/ KNI-272 complex with a cyclic urea class of inhibitor. (**A**) the H-bond interactions of the bridging water molecule at the interface of HIV-1 protease/ KNI-272 complex (PDB code: 1hpx; K_i: 5.5 pM). The bridging water molecule forms four ordered H-bonds with the HIV-1 protease/ KNI-272 complex. (**B**) After the displacement of the bridging water molecule by a cyclic urea inhibitor (PDB code: 1ajx; K_i: 12.2 nM), the modified ligand forms two ordered H-bonds at the water binding site. Thus, this water displacement process reduces two ordered H-bonds. This example is merely used to demonstrate that displacing a bridging water molecule into bulk water changes the number of ordered H-bonds. Because the cyclic urea inhibitor in (B) is entirely different from KNI-272 in (A), the difference in binding affinity between the two inhibitors has no relationship with the change in binding affinity coming from water displacement.

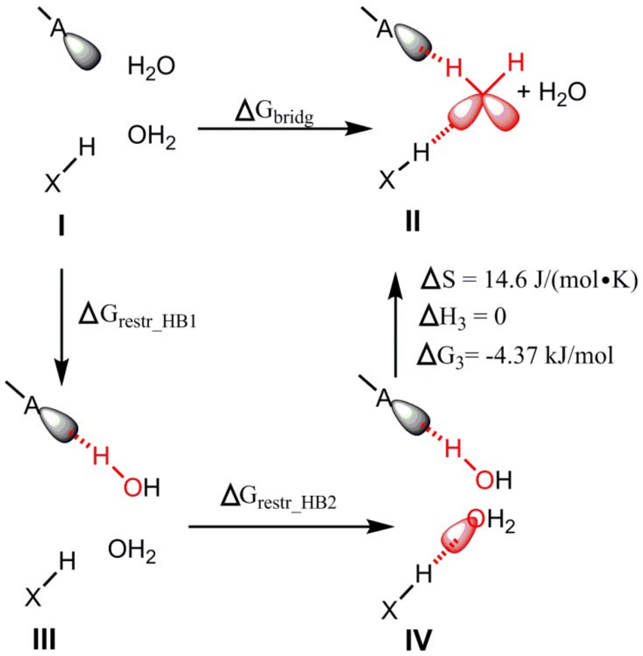

Figure S2. Thermodynamic cycle for deriving the free energy change (ΔG) for forming a bridging water molecule that has optimal interactions with two polar atoms. Assume the bridging water molecule forms two H-bonds with polar atoms A and H of H-X. The ΔG for forming the bridging water molecule from bulk water (ΔG_{bridg}) is the ΔG of the process from I to II. This cycle indicates that the ΔG_{bridg} is the sum of the ΔGs for restricting the H-bonds between the polar atoms and bulk water to their optimal geometries (from I to III and III to IV) and the ΔG for the formation of the bridging water molecule from IV. In the process from IV to II, one of the two water molecules in IV becomes the bridging molecule by displacing the other water molecule into bulk water. The enthalpy change on going from IV to II is zero because the strength of the broken H-bond equals to that of the formed H-bond. The net ΔS on going from IV to II is 14.6 J/(mol·K) (Figure S3), which corresponds to a favorable ΔG of –4.37 kJ/mol at 298 K. If the polar atoms are from flexible protein side-chains or are flexible ligand atoms, they become less flexible once they interact with the bridging water. Assume the ΔGs for restricting an unrestricted H-bond and a flexible polar atom are ΔG_{restr_HB} and $\Delta G_{restr\_polar\_atom}$ respectively, the ΔG_{bridg} for the bridging water molecule with two restricted H-bonds is:

$$\Delta G_{bridg} = \Sigma\Delta G_{restr\_HB} + \Sigma\Delta G_{restr\_polar\_atom} - 4.37 \text{ (kJ/mol)} \tag{S3}$$

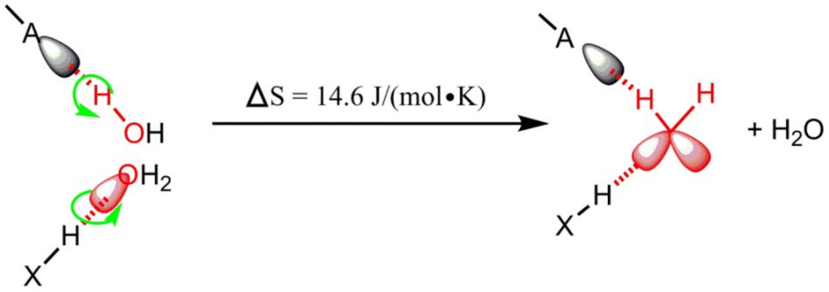

Figure S3: Entropy change for the formation of a bridging water molecule from two water molecules that have an ordered H-bond. The entropy difference between a restrained water molecule in ice and an unrestrained water molecule is 22.0J/mol, which results from the increment of the translational and orientational degrees of freedom of the water molecule. The difference between the numbers of translational and orientational degrees of freedom between a bulk water molecule and a water molecule in ice is 6. The water molecules on the left side, which have an ordered H-bond, have only one orientational degree of freedom. Thus, the entropy change for the water molecule that becomes the bridging water molecule is -22.0*(1/6) = –3.7 J/(mol·K). The entropy change for the water that is released to bulk water is 22.0*(5/6) = 18.3 J/(mol·K). Thus the net entropy change for the formation of a bridging water molecule from two water molecules that have an ordered H-bond is 14.6 J/(mol·K).

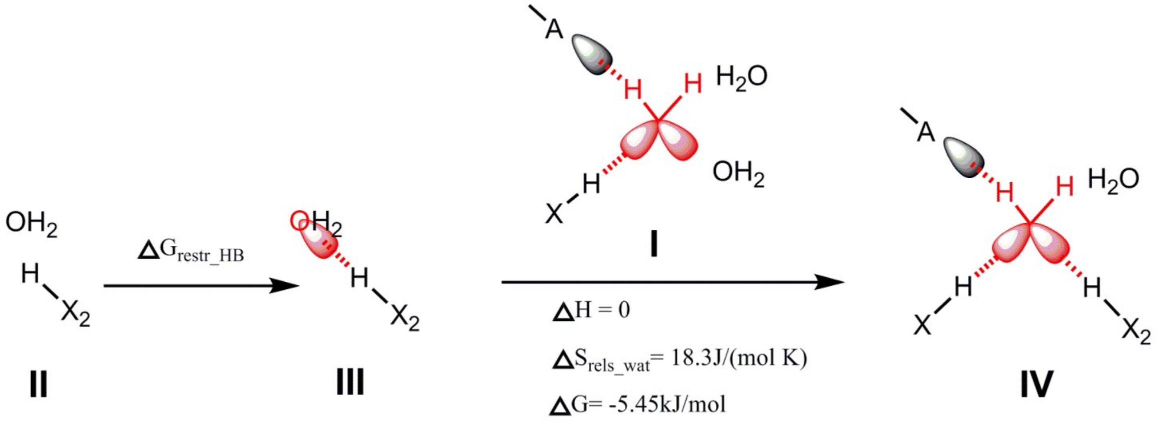

Figure S4. Free energy change for forming an ordered H-bond between a bridging water molecule and a polar atom from protein-ligand complex. Assume the bridging water molecule is the water molecule colored red in **I** and the polar atom from protein-ligand complex is the hydrogen atom of H-X_2 in **II**. First, interaction between H-X_2 and bulk water is restricted to its optimal geometry (II to III), which has a free energy change of ΔG_{restr_HB}. If the polar atom is flexible, free energy is required to restrict the polar atom ($\Delta G_{restr\_polar\_atom}$). Then the bridging water molecule in **I** displaces the restricted water molecule in **III** so that the bridging water molecule forms an ordered H-bond with H-X_2 (from **III** to **IV**) and the restricted water molecule in **III** is released to bulk water. The enthalpy change and entropy change of this process is 0 and 18.3 J/(mol.K) respectively, indicating the free energy change of the process from **III** to **IV** is -5.45 kJ/mol at 298K. Thus, free energy change for forming an ordered H-bond between a bridging water molecule and a polar atom from protein-ligand complex ($\Delta\Delta G_{fix\_wat}$) is $\Delta G_{restr\_HB} + \Delta G_{restr\_polar\_atom}$ -5.45kJ/mol.

Because the ΔG_{fix_wat} for a bridging water molecule with two restrained H-bonds is $\Sigma\Delta G_{restr\_HB}$

$+\Sigma\Delta G_{restr\_polar\_atom}$ - 4.37 (kJ/mol), the equation of the ΔG_{fix_wat} for the bridging water molecule that interacts with n ($4 \geq n \geq 2$) polar atoms and has (4-n) H-bond interactions with bulk water is:

$$\Delta G_{fix\_wat} = \Sigma\Delta G_{restr\_HB} + \Sigma\Delta G_{restr\_polar\_atom} - (n-1)*5.45 + 1.08 \qquad \text{(S4)}$$

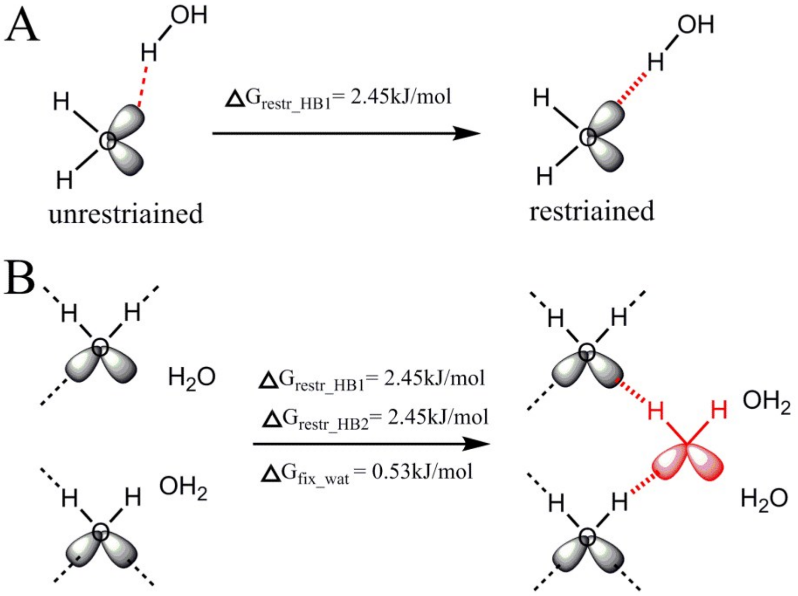

Figure S5: Free energy change (ΔG) of the water freezing process. (A) The ΔG for restricting a $H_2O \ldots HOH$ H-bond. For the process from unrestrained H-bond to well-oriented restrained H-bond, the H-bond becomes stronger ($\Delta H < 0$), but the water molecule's translational and orientational degrees of freedom decrease ($\Delta S < 0$). For the process from bulk water to ice at 273.2K, ΔG, ΔH and ΔS from water to ice are 0, -6.01kJ/mol and -22.0J/(mol.K) respectively. As one mole of water molecules in ice contains two moles of hydrogen bonds, ΔH of process for restricting a $H_2O \ldots HOH$ H-bond as shown in this Figure is -3.00 kJ mol^{-1}. As the restrained water molecule shown in this Figure can rotate in one direction, five of six degrees of freedom are restricted; the ΔS for this process is -18.3 J/(mol.K). Thus, at 298K, the ΔG for restricting a $H_2O \ldots HOH$ H-bond in bulk water (ΔG_{restr_HB}) is: -3.00 -298 *(-18.3)/1000 = 2.45 kJ/ mol. **(B)** The bulk water freezing process and theΔG of the process. This Figure shows that a water molecule from bulk water becomes a bridging water molecule by forming ordered H-bonds with the lone pair electron of one fixed water molecule and the hydrogen atom of another fixed water molecule. The process is a bulk water freezing process because bulk water would convert to ice by repeating this process. The ΔG of this process can be obtained from the equation for calculating ΔG of the process that a bridging water molecule with two ordered H-bonds is formed from bulk water, which is equation (2) of the manuscript:

$$\Delta G_{fix\_wat} = \Sigma\Delta G_{restr\_HB} + \Sigma\Delta G_{restr\_polar\_atom} - 4.37 \text{ (kJ/mol)} \quad \text{(S5)}$$

$\Sigma\Delta G_{restr\_polar\_atom}$ is the sum of the free energy for restricting flexible polar atoms, which is zero for the bulk water freezing process. $\Sigma\Delta G_{restr\_HB}$ is the sum of the ΔG for restricting two H_2O…HOH H-bonds. As the ΔG for restricting a H_2O…HOH H-bond is 2.45 kJ/mol at 298 K, the ΔG of the water freezing process is 0.53 kJ/mol at 298 K.

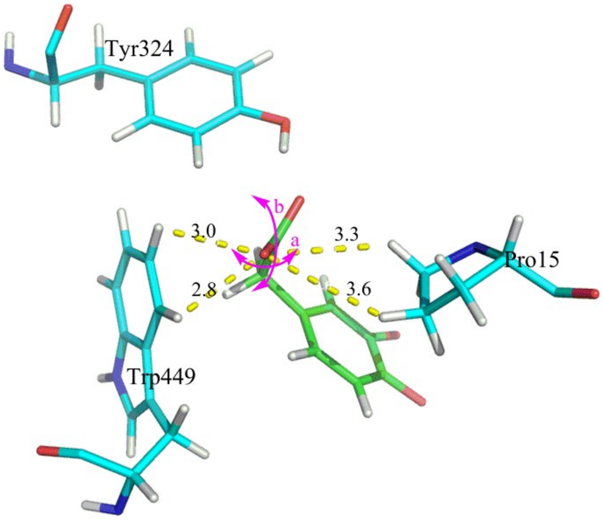

Figure S6: The flexibility of the carboxyl group from the inhibitor of 3PCN. Trp449 and Pro15 are the residues close to the carboxyl group from the inhibitor of 3PCN and reduce the flexibility of the carboxyl group. The distances between the oxygen atoms of the carboxyl group and the close atoms from residues are larger than the sum of Van der Waals radii of the atoms (~2.6Ǻ). There are two rotatable bonds between the carboxyl group and the fixed benzene ring. Thus the carboxyl group can rotate around the bonds as shown with “a” and “b”. Moreover, the side chains of Trp449 and Tyr324 are flexible. The binding site is also flexible (otherwise, the inhibitors would not be able to enter the binding site). Thus, the carboxyl group is flexible.

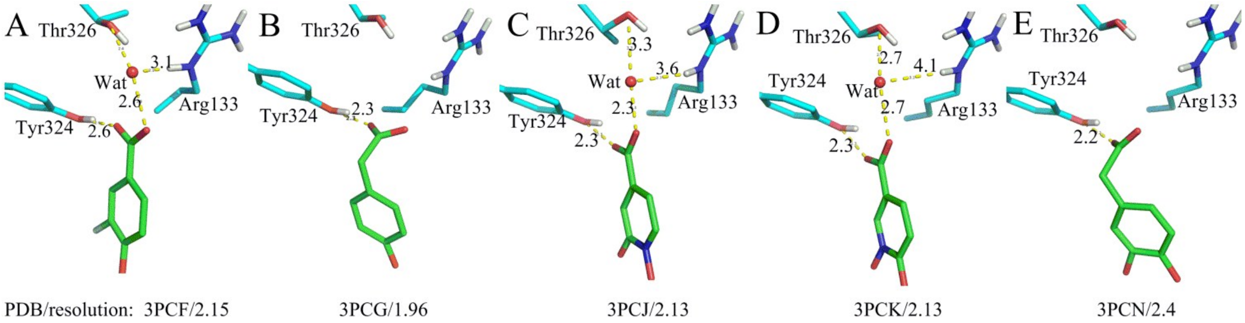

Figure S7. Flexibilities of the polar atoms interacting with a bridging water molecule have large effect on whether the bridging water can be observed in the crystal structure. This figure compares the water-binding sites located at interfaces of protocatechuate 3,4-dioxygenase and its inhibitors. Bridging water molecules are observed in 3PCF (A), 3PCJ (C) and in 3PCK (D), but are not observed in 3PCG (B) and 3PCN (E). The carboxyl groups of the inhibitors in 3PCF, 3PCJ and in 3PCK are not flexible because the carboxyl groups are conjugated with the fixed aromatic rings. But the carboxyl groups of the inhibitors in 3PCG and 3PCN are flexible because of the presence of sp^3 carbons between the carboxyl groups and the fixed aromatic rings. Thus, bridging water molecules are not observed in 3PCG and 3PCN because the carboxyl groups of the inhibitors are flexible. In this example, the resolutions of the crystal structures have no relationship with bridging water molecules because 3PCG has highest resolution and the bridging water molecule is not observed in its water binding site.

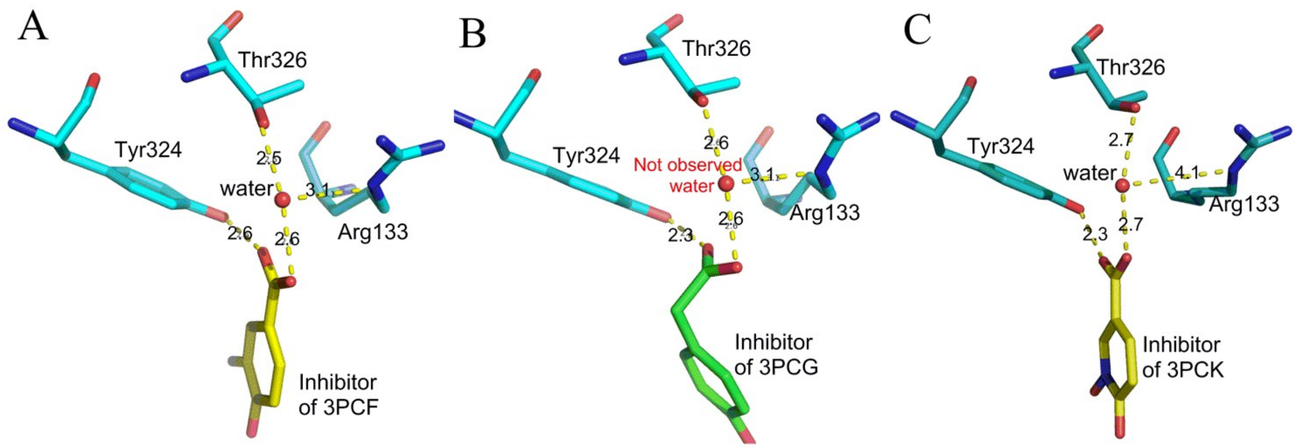

Figure S8: Little effect of the H-bond strength of a bridging water molecule on the free energy change for forming the bridging water molecule from bulk water. The data are the distances between the heavy atoms. The water binding sites between protocatechuate 3,4-dioxygenase and its three inhibitors are shown in (A) (pdb code: 3PCF), (B) (pdb code: 3PCG), and (C) (pdb code: 3PCK). Bridging water molecules are observed in water binding sites of 3PCF and 3PCK, but not in 3PCG. The bridging water molecule shown in (B) is obtained by superimposed the structure of 3PCG to the structure of 3PCF (see Methods). The bridging water molecules in (A) and (B) interact with an oxygen atom of the inhibitor carboxyl group, the hydroxyl group of Thr326, and the polar hydrogen atom of Arg133. The bridging water molecule in (C) interacts only with an oxygen atom of the inhibitor carboxyl group and the hydroxyl group of Thr326, but not with the polar hydrogen atom of Arg133 because of the large distance between the water oxygen atom and nitrogen atom of Arg133. Because the H-bond interaction between positively charged hydrogen of Arg with water is very strong, the bridging water molecule in (B) has much stronger interactions than the bridging water molecule in (C). However, the bridging water molecule in (B) is not observed, while the bridging water molecule in (C) is observed. Thus, the strong H-bond between a bridging water molecule and a flexible polar atom has little effect on the free energy change for forming the bridging water molecule from bulk water. As will be demonstrated in main manuscript, the water molecule is not observed in 3PCG because the carboxyl group is flexible.

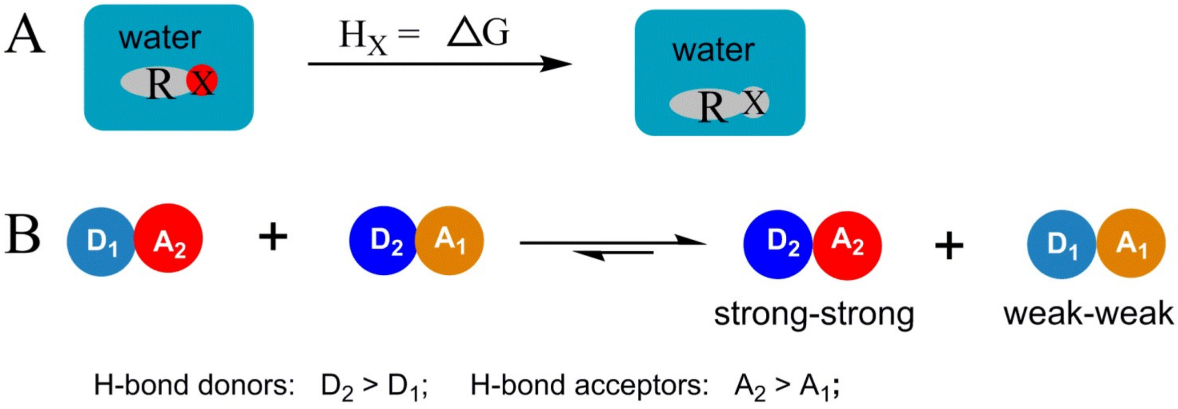

Figure S9: H-bond pairing principle and the free energy changes of H-bond pairing processes: (**A**) The definition of the H-bonding capability. The H-bonding capability of an atom is defined as the free energy change (ΔG) relevant to noncovalent interactions of the process in which an atom is depolarized in water (X colored red represents polar atoms; X colored grey represents apolar atoms). Atoms with higher H-bonding capabilities can form stronger H-bonds with water. (**B**) The H-bond competing pairing process between two H-bond donors (D_1 and D_2) and two H-bond acceptors (A_1 and A_2). The H-bonding capabilities for D_1, D_2, A_1, and A_2 are H_{D1}, H_{D2}, H_{A1}, and H_{A2} respectively. The free energy change of the H-bond pairing process shown in Figure (B) is[1]

$$\Delta G = -(H_{D2} - H_{D1}) \times (H_{A2} - H_{A1})/k \tag{S6}$$

where k is a positive constant and equals to the H-bond capability of the hydrogen atom or the lone pair electron of water. We assume D_2 is a stronger H-bond donor than D_1 ($H_{D2} > H_{D1}$) and A_2 is a stronger H-bond acceptor than A_1($H_{A2} > H_{A1}$). The free energy change of the process in this figure is less than 0. Thus, the H-bond pairing process favors the strong-strong H-bond and weak-weak H-bond pairings. If a protein-ligand H-bond is strong-strong pairing, increasing the H-bonding capability of the ligand atom can enhance the contribution of the H-bond to ligand binding affinity. Similarly, if a protein-ligand H-bond is weak-weak pairing, decreasing the H-bonding capability of the ligand atom can enhance the contribution of the H-bond to ligand binding affinity.

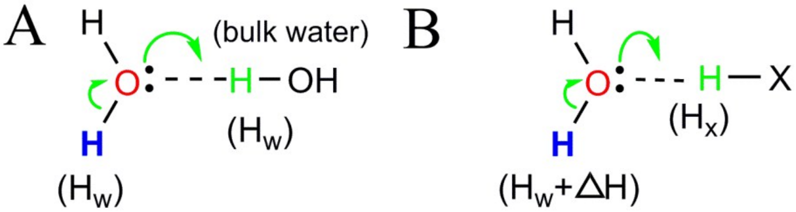

Figure S10: Effect of H-bond cooperativity on the H-bonding capability of bridging water molecules. (**A/B**) The H-bonds between the oxygen atoms (colored red) and the hydrogen atoms (colored green) increase the positive charges of the hydrogen atoms colored blue. The H-bonding capabilities calculated based on the approaches in Methods are defaulted for the environment of aqueous solution. The H-bonding capability for the hydrogen atom of bulk water is H_w, which is based on the prerequisite that the oxygen atom form H-bonds with bulk water. If the O...H-X H-bond interaction in (B) is different from the H-bond interaction between water molecules in (A), the H-bonding capability of the hydrogen atom colored blue in (B) should be adjusted because of H-bond cooperativity[5-7]. If the hydrogen atom in H-X is more positively charged than the hydrogen atom of water, the hydrogen atom colored blue in (B) is also more positively charged than the blue colored hydrogen atom in (A). Thus, if the H-bonding capability of the hydrogen atom in H-X (H_X) is larger than that of water (H_W), the H-bonding capability of the water hydrogen atom colored blue in (B) is larger than H_W ($\Delta H > 0$). Similarly, if the H-bonding capability of the hydrogen atom in H-X (H_X) is lower than that of water (H_W), the H-bonding capability of the water hydrogen atom colored blue in (B) is lower than H_W ($\Delta H < 0$).

References

1. Chen, D.; Oezguen, N.; Urvil, P.; Ferguson, C.; Dann, S. M.; Savidge, T. C., Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv.* **2016**, *2*, e1501240.
2. Orville, A. M.; Lipscomb, J. D.; Ohlendorf, D. H., Crystal structures of substrate and substrate analog complexes of protocatechuate 3,4-dioxygenase: endogenous Fe3+ ligand displacement in response to substrate binding. *Biochemistry* **1997**, *36*, 10052-66.
3. Elgren, T. E.; Orville, A. M.; Kelly, K. A.; Lipscomb, J. D.; Ohlendorf, D. H.; Que, L., Jr., Crystal structure and resonance Raman studies of protocatechuate 3,4-dioxygenase complexed with 3,4-dihydroxyphenylacetate. *Biochemistry* **1997**, *36*, 11504-13.
4. Orville, A. M.; Elango, N.; Lipscomb, J. D.; Ohlendorf, D. H., Structures of competitive inhibitor complexes of protocatechuate 3,4-dioxygenase: multiple exogenous ligand binding orientations within the active site. *Biochemistry* **1997**, *36*, 10039-51.
5. Blanco, S.; Pinacho, P.; Lopez, J. C., Hydrogen-Bond Cooperativity in Formamide2 -Water: A Model for Water-Mediated Interactions. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 9331-5.
6. Lei, Y.; Li, H.; Zhang, R.; Han, S., Theoretical study of cooperativity in biotin. *J. Phys. Chem. B.* **2007**, *111*, 14370-7.
7. Stokely, K.; Mazza, M. G.; Stanley, H. E.; Franzese, G., Effect of hydrogen bond cooperativity on the behavior of water. *Proc. Natl. Acad. Sci. U S A* **2010**, *107*, 1301-6.