

# SCIENTIFIC REPORTS



OPEN

## The origin of the cooperativity in the streptavidin-biotin system: A computational investigation through molecular dynamics simulations

Received: 02 February 2016

Accepted: 13 May 2016

Published: 01 June 2016

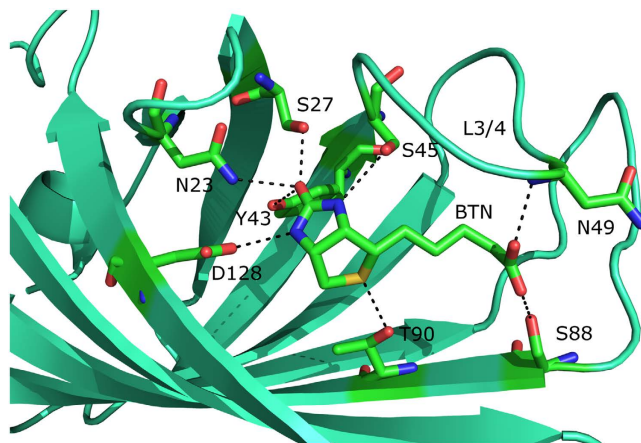
Fengjiao Liu<sup>1</sup>, John Z. H. Zhang<sup>2,3</sup> & Ye Mei<sup>1,3</sup>

Previous experimental study measuring the binding affinities of biotin to the wild type streptavidin (WT) and three mutants (S45A, D128A and S45A/D128A double mutant) has shown that the loss of binding affinity from the double mutation is larger than the direct sum of those from two single mutations. The origin of this cooperativity has been investigated in this work through molecular dynamics simulations and the end-state free energy method using the polarized protein-specific charge. The results show that this cooperativity comes from both the enthalpy and entropy contributions. The former contribution mainly comes from the alternations of solvation free energy. Decomposition analysis shows that the mutated residues nearly have no contributions to the cooperativity. Instead, N49 and S88, which are located at the entry of the binding pocket and interact with the carboxyl group of biotin, make the dominant contribution among all the residues in the first binding shell around biotin.

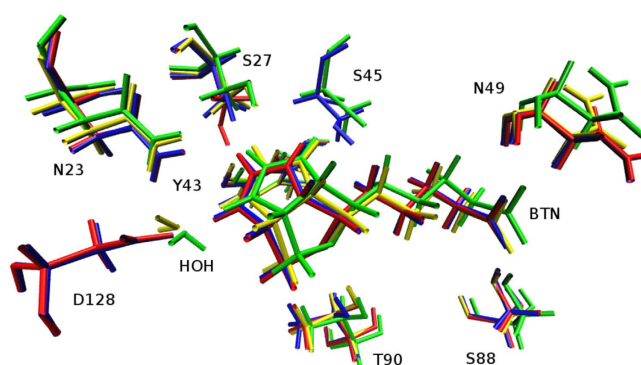
Cooperativity of residues in modulating the affinity of protein-ligand binding has been frequently used by experimental chemists<sup>1</sup>, despite that this concept lacks of a solid theoretical definition<sup>2–4</sup>. Through double-mutant cycles, the excess loss of binding affinity in double mutant relative to the sum of that for two single mutants has been considered as a direct evidence of the cooperativity between these two mutated residues<sup>5</sup>. It is believed that neutral residues within 7 Å or charged residues within 10 Å in the receptor may have significant cooperativity in modulating its binding affinity to ligands<sup>5,6</sup>.

Streptavidin-biotin complex is one of the strongest binding partners occurring in nature and is far beyond the normal protein-ligand binding strength<sup>7–9</sup> and has been widely used in biochemical sensing applications, especially in the identification of possible new drug targets<sup>10</sup>. As shown in Fig. 1, biotin is confined in the active site of streptavidin by eight hydrogen bonds, as well as van der Waals interaction among non-polar groups<sup>11</sup>. Among these “first shell” residues, five of them bind to the biotin ureido group and are deeply buried inside the pocket. The side chains of Asn23, Ser27, and Tyr43 form hydrogen bonds with the ureido oxygen atom. Ser45 and Asp128 interact with the ureido H atoms from two sides. Previous molecular dynamics simulations and energy decomposition studies<sup>12</sup> suggested that the dominant contribution to this strong binding affinity was from van der Waals interaction between the ligand and the non-polar groups, especially the tryptophan residues, encompassing the ligand. Strong electrostatic and hydrogen bond interactions between the ligand and the receptor are largely canceled out by the penalty for removing water molecules from the binding site and the desolvation of the ligand. However, some recent studies have identified the significant contribution from electronic polarization effect, which had not been taken into account explicitly in Kollman’s work<sup>13,14</sup>. Besides, some residues that bind to the “first shell” residues but not to biotin may also play an important role in modulating the binding affinity through the hydrogen bond cooperativity<sup>14,13</sup>. In addition, a flexible loop (Ser45 to Ser52) connecting  $\beta$  strands 3 and 4

<sup>1</sup>State Key Laboratory of Precision Spectroscopy, School of Physics and Materials Science, East China Normal University, Shanghai 200062, China. <sup>2</sup>Department of Physics, School of Physics and Materials Science, East China Normal University, Shanghai 200062, China. <sup>3</sup>NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China. Correspondence and requests for materials should be addressed to Y.M. (email: ymei@phy.ecnu.edu.cn)



**Figure 1. Binding of biotin to the wild-type streptavidin.** The key hydrogen bonds are marked as dashed lines.



**Figure 2. Superposition of the wild-type (blue), S45A (red), D128A (green) and S45A/D128A (yellow) biotin complexes in the biotin binding site.**

(L3/4)<sup>15,16</sup>, facilitates ligand binding by excluding solvent molecules from the active site, enhancing the direct hydrogen-bonding interactions and hydrophobic interactions with the loop closed. It could cause 27.1 kcal/mol free energy increase from the loop's open state to closed state, which is estimated by General and Meriovitch using the hypothetical scanning molecular dynamics (HMSD) method<sup>17,18</sup>.

Some studies focusing on the hydrogen bond network show that the streptavidin biotin hydrogen bonds indeed display exceptionally large free energy alterations after mutation<sup>19–22</sup>. Freitag *et al.* pointed out that D128 acts as an intermediate on the way of biotin dissociation by a series of site-directed mutagenesis, biophysical, and X-ray crystallographic approaches<sup>23</sup>. A number of studies on S45A mutant indicated that the S45A mutation exhibits a 1,700-fold greater dissociation rate and 907-fold lower equilibrium affinity for biotin relative to the wild-type streptavidin at 37 °C, exhibiting its important role in modulating both the equilibrium and transition state energetics<sup>22</sup>. Isothermal titrating calorimetry (ITC)<sup>24</sup> experiments performed on these mutants by Hyre *et al.* indicated that the double-mutant resulted in 10.1 kcal/mol loss of binding affinity, while that caused by mutation of S45A and D128A single-mutations were just 4.3 and 4.4 kcal/mol, respectively. This result demonstrated that the loss of binding affinity in the S45A/D128A double-mutant (DM) was not just a linear combination of single-mutant energetic perturbations but showed a cooperative effect. They suggested that this effect is caused by the structural rearrangements at the S45 position when the D128 carboxyl group is removed, which masks the true energetic contribution of the D128-biotin interaction<sup>11</sup>. However, the structural differences among the wild-type and the mutants are very small, which can be seen in Fig. 2 and qualitatively from Table 1. These data indicate that the cooperativity does not come from the global structure rearrangement after these mutations.

Although crystallographic and calorimetric experiments have proposed a conjecture that these pivotal hydrogen bonds act cooperatively, direct evidence from computer simulation has not yet been presented. A large number of studies have shown that with proper parameterization molecular mechanics (MM) is quite successful in modeling biological molecules, such as proteins and nucleic acids<sup>25–30</sup>. Nevertheless, in standard force fields, e.g., CHARMM<sup>26</sup> and AMBER<sup>31–34</sup>, the atomic charge magnitudes do not change in response to the chemical environment. Therefore, they only give a mean-field representation of the charge distribution, and fail to capture the adaptation of electron density to the environment, which is highly inhomogeneous and protein-specific. Electronic polarization is critical for an accurate description of electrostatic interaction, due that the electrostatic

	RMSD (Å)
S45A	0.13
D128A	0.55
DM	0.28

**Table 1.** The root mean square deviation of the backbone atoms in the mutants from WT.

interaction depends not only on their distance of atoms but also on their locations in the protein and their local solvent environment<sup>35–38</sup>. Cooperativity is a typical nonadditive phenomenon, which poses a challenge to traditional force field. Ji *et al.* proposed the polarized protein-specific charge (PPC) scheme<sup>38</sup> that can reflect the specificity of charge distribution in a certain protein, based on a quantum mechanical fragmentation method termed the molecular fractionation with conjugate caps. This charge scheme is further refined by Zeng *et al.* via improving its numerical stability<sup>39</sup>. The viability and effectiveness of this charge model have been proved in several studies<sup>40</sup>, including biotin and its analogues binding to the wild type (strept) avidin and the mutant<sup>14,41–43</sup>. Meanwhile, quantum mechanical calculation has been applied to this system by Houk *et al.*, and they found that these hydrogen bonds act cooperatively<sup>13</sup>. The charged D128 residue provides the driving force for the cooperativity in the hydrogen-bonding network by greatly polarizing the ureido of biotin. If the D128 residue is mutated, the hydrogen bond network connecting the “first shell” and the “second shell” residues is disrupted. Unfortunately, due that QM calculation is too costly for the whole protein, the thermal effect had not been taken into consideration in their study.

When structure ensemble is available, the free energy can be estimated using various methods differ by their rigor and speed. All these methods can be roughly classified into two categories, i.e. the pathway method and the end-state method<sup>44</sup>. As suggested by the name, pathway methods compute the sum of small changes along a geometrical or alchemical route connecting the initial and the final states, while end-state methods consider the conformations of the free and the bound states only and compute the binding affinity by taking a difference. No intermediate states are considered in end-state methods. The thermodynamic perturbation (TP)<sup>45</sup>, thermodynamic integration (TI)<sup>46–48</sup> and Bennett acceptance ratio (BAR)<sup>49,50</sup> methods are the typical representatives of pathway methods. The end-state methods usually refer to the linear interaction energy (LIE)<sup>51,52</sup> and the molecular mechanics/Poisson-Boltzmann with surface area (MM/PBSA) method<sup>53–55</sup>, along with its variant utilizing the generalized Born solvation model (MM/GBSA). Although the pathway methods can provide results within 1 to 2 kcal/mol from the experimental measurement, their computational expenses are usually considerably higher than that of the end-state methods<sup>44</sup>. The LIE method requires a training process and is more suitable for the calculation of the (relative) binding free energies of a series of ligands<sup>52</sup>. The (strept)avidin-biotin system has been widely studied using several methods, including FEP<sup>56</sup>, LIE<sup>57</sup>, and MM/PBSA<sup>58–61</sup>. All of the computational results showed good correlation with the experimental values<sup>62</sup>. Therefore, the MM/PBSA method was utilized in this work for the calculation of the binding affinities.

In this work, classical molecular dynamics (MD) simulations have been carried out to study the wild-type streptavidin and its D128A, S45A and S45A/D128A mutants complexed with biotin. The electrostatic polarization effect is taken into account by utilizing the PPC scheme. The binding free energies are estimated using the MM/PBSA method. The results show that this cooperativity comes from both the enthalpy and entropy contributions. And the former contribution mainly comes from the alternations of solvation free energy.

## Methodology

**Molecular dynamics simulation.** The initial structures for the wild-type, S45A, D128A and S45A/D128A streptavidin, which will be denoted as WT, S45A, D128A and DM hereafter, were downloaded from Protein Data Bank (entry: 1MK5, 1DF8, 1SWT and 1MEP, respectively). Streptavidin is a dimer-of-dimer, and the complete structures were generated by applying the symmetry operation with Discovery Studio Visualizer except for DM, of which the coordinates for all the four units are available in the crystal structure. The coordinates of the missing residues are predicted by the ModLoop server<sup>63,64</sup>. AMBER03 force field<sup>65</sup> and the general AMBER force field (GAFF)<sup>66</sup> were applied to the protein and ligand, respectively. All hydrogen atoms in protein were automatically added by the LEaP module in AMBER package. Each complex was soaked in a periodic box of TIP3P water<sup>67</sup>, with the minimal distance between the complex and the boundary of the box set to 12 Å. Sodium ions were added to neutralize the charge. Crystal water molecules were reserved when preparing the system. Each system was first optimized with weak restraint exerted on the ligand molecules and the proteins (excluding those missing residues optimized by the ModLoop server) and then was fully relaxed until the energy gradient was below  $1.0 \times 10^{-4}$  kcal/(mol Å). With this energy-minimized structure, the PPCs were fitted using the molecular fractionation with conjugate caps/Poisson-Boltzmann (MFCC/PB) method<sup>38,39</sup>. The quantum mechanical calculations were carried out at B3LYP/6-31G(d) level using Gaussian 09<sup>68</sup>. The calculated atomic charges were then averaged over the four chains and were used in the subsequent MD simulations. Parameters other than the atomic charges were kept intact. Each system was gradually heated up to 300 K in 100 ps, and then further relaxed for 20 ns at NPT ensemble. The temperature was regulated with the Langevin dynamics heat coupling scheme and the collision frequency was set to  $2.0 \text{ ps}^{-1}$ . The pressure relaxation time was set to 1.0 ps. SHAKE algorithm was applied to restrain all covalent bonds involving hydrogen atoms<sup>69</sup>. The integral time step was set to 2 fs. The van der Waals (vdW) interaction was truncated at 10 Å. Long range Coulomb interaction was calculated using particle mesh Ewald (PME) with 10 Å cutoff in real space<sup>70</sup>.

**MM/PBSA calculation of the binding free energies.** The binding affinity was calculated by<sup>44</sup>

$$\Delta G_{bind} = G(RL) - G(R) - G(L), \quad (1)$$

where  $R$  and  $L$  stand for the receptor and the ligand. The free energy of each reactant was estimated as

$$\begin{aligned} G &= H - TS \\ &= E + \Delta G_{pol} + \Delta G_{npol} - TS, \end{aligned} \quad (2)$$

with

$$E = E_{bonded} + E_{es} + E_{vdw}, \quad (3)$$

where  $E_{bonded}$ ,  $E_{es}$  and  $E_{vdw}$  are the bonded energy, the Coulomb interaction energy and the vdW interaction energy, respectively.  $\Delta G_{pol}$  is the polar solvation free energy, which was calculated by solving the Poisson-Boltzmann equation.  $\Delta G_{npol}$  is the nonpolar contribution to the solvation free energy, which was approximately estimated as a function of the solvent-accessible surface area (SASA) as  $G_{npol} = \gamma A_s + b$  with  $\gamma = 0.00542 \text{ kcal}/(\text{mol} \cdot \text{\AA}^2)$  and  $b = 0.92 \text{ kcal/mol}$ . The surface area  $A_s$  was calculated with the molsurf package<sup>71</sup>.  $S$  is the entropy contribution estimated using the normal-mode approximation. The binding free energy is finally expressed as

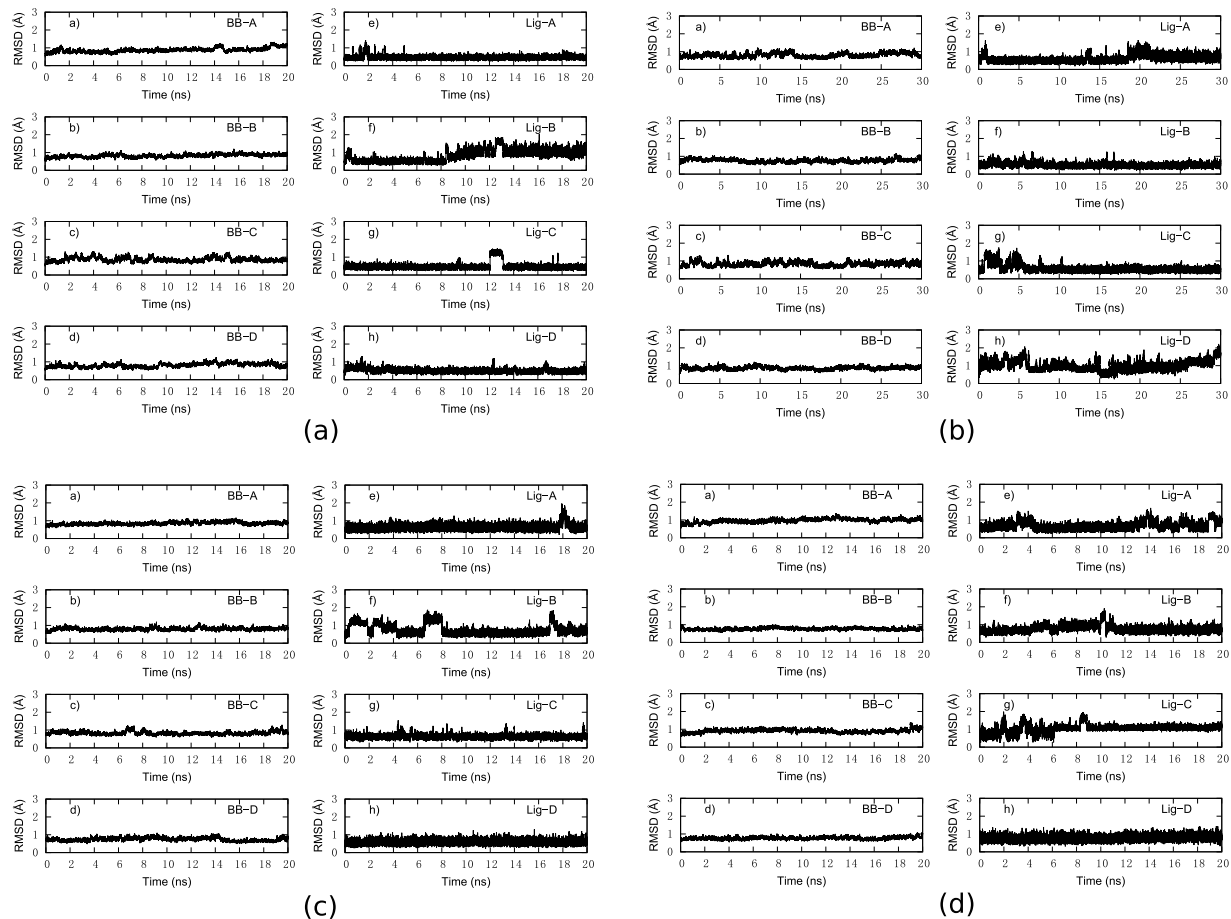
$$\begin{aligned} \Delta G_{bind} &= \Delta H - T\Delta S \\ &= \Delta E_{bonded} + \Delta E_{es} + \Delta E_{vdw} + \Delta\Delta G_{pol} + \Delta\Delta G_{npol} - T\Delta S. \end{aligned} \quad (4)$$

A 20-ns production MD simulation was performed for each system. However, the hydrogen bonding network in chain A of S45A was not as stable as those in other chains after about 17 ns. Therefore, the last 3-ns trajectory was abandoned for the S45A mutant. All of the terms in Eq. 4 were averaged over the snapshots extracted from the most stable MD trajectory, i.e. 4–20 ns for WT, D128A and DM, and 4–17 ns for S45A. Altogether, eight hundred snapshots were used for the calculation of enthalpy. The “mbondi2” radii were used in the PB calculations, and the grid spacing was set to 0.5 Å. The exterior and interior dielectric constants were set to 80 and 1, respectively. One hundred snapshots were used for the entropy calculation using the Nmode program. All the molecular dynamics simulations and MM/PBSA analysis were conducted with Amber12 and AmberTools package.

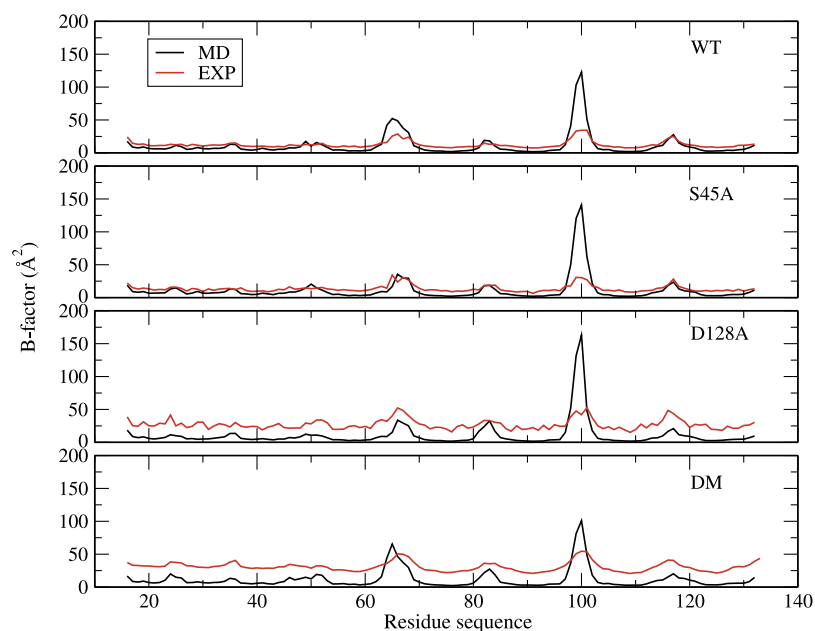
## Result and Discussion

**Structure stability.** The overall structure variation of the protein and ligand molecules during the whole simulation is shown in Fig. 3. The backbone RMSDs of the proteins are always around 1 Å with very small fluctuations, indicating remarkable stabilities of the protein scaffolds. Besides, significant variation of conformation or orientation for ligand in the binding pocket of the four system has not been detected. The RMSDs of biotin in the binding pocket are marginally larger than those of the protein backbone, but are still below 2 Å in majority of the simulation time. Due to the slightly different conformations of the four chains and the random velocities assigned to the atoms at the beginning of the simulations, notable differences among the four chains can be observed, especially for the ligand RMSDs. The simulated and experimental B-factors are shown in Fig. 4. They only qualitatively agree with each other. This is not unexpected, because the fluctuations of the proteins in crystal are significantly impeded by packing interaction of neighboring units. Nonetheless, the flexible regions can be detected from both the simulated and experimental data. The crystal water molecules are retained when building the systems. Some of them are located in the biotin binding pocket and are close to residue 128. However, these water molecules are not that stable during the simulations of D128A and DM. Frequent exchange with surrounding water molecules can be observed. Furthermore, no significant change in the internal structure of the protein can be detected, which can be read from the contact map shown in Fig. 5 defined as the average distance between  $C_\alpha$  atoms. Compared with WT, S45A leads to only small alternations of the distance between nonterminal  $C_\alpha$  atoms, which are generally limited within  $\pm 1$  Å. Due to the large flexibility of the C-terminal residues, they show considerable change of the distances to all the other residues. However, it is believed that these terminal residues have little impact on the binding affinity of the ligand, because their interaction with the ligand are strongly screened by solvent molecules. Relatively large difference can be seen for Q24 and Y83 in D128A and DM. These residues are located in the loop regions of the protein, which show great flexibility as indicated by the B-factors shown in Fig. 4.

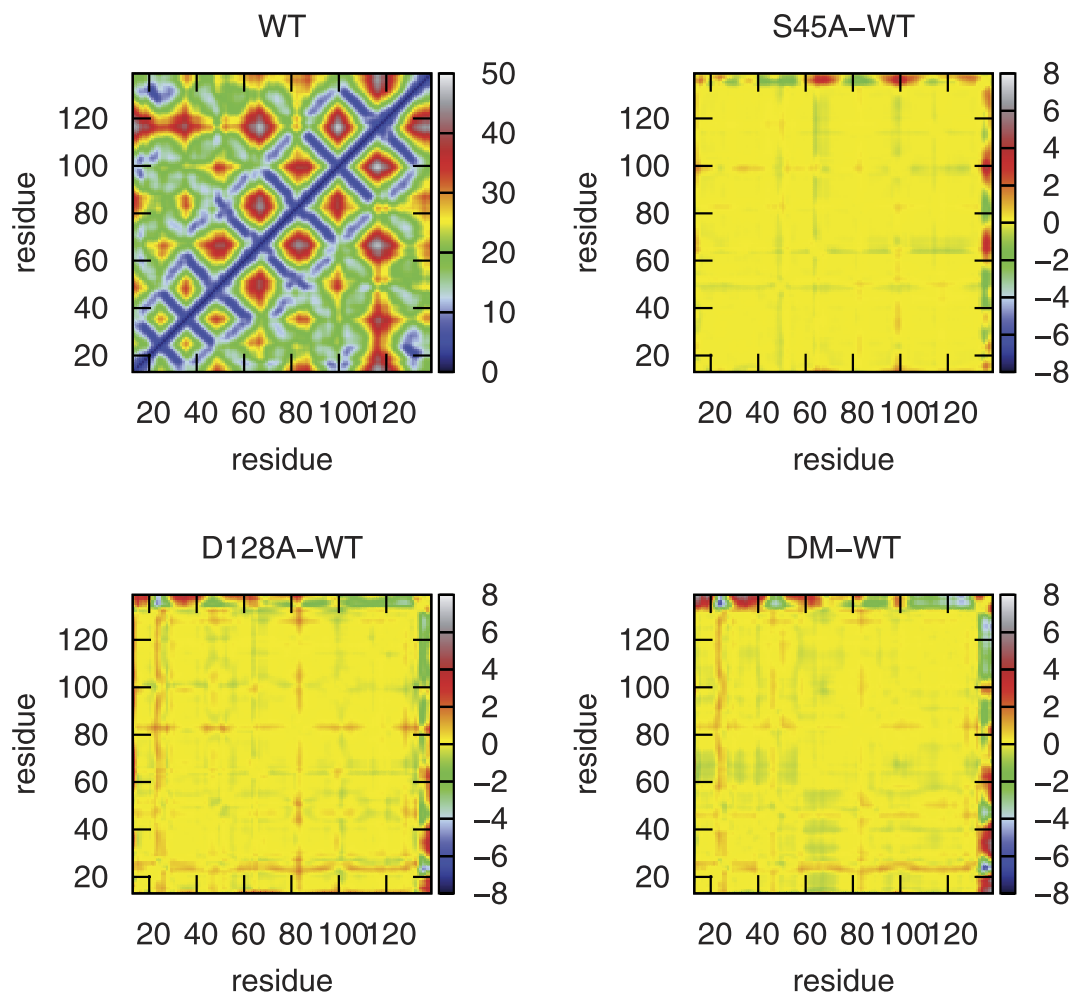
Interaction of the hydrogen bonds between streptavidin and biotin is essential for the binding affinity, of which the strength can be indirectly inspected by the variation of bond length as shown in Fig. 6. The carboxyl group of biotin is partially exposed to the solvent. Invasion of water molecules destabilizes the hydrogen bonds between this carboxyl group and the proteins (N49 and S88). Therefore, these two hydrogen bonds are interrupted and reconstructed repeatedly during the simulations. In WT, all other hydrogen bonds between the ligand and the protein are extremely stable and are consistent among all the chains. The largest fluctuation is seen for the hydrogen bond connecting the ligand and N23. N23 is anchored near biotin by the hydrogen bond between N23 and D128 (see Fig. 7), as well as the direct hydrogen bonding interaction between N23 and the ligand. When D128 mutates into alanine as in D128A and DM, the side chain of N23 can no longer maintain this orientation and its hydrogen bond to biotin broke quickly at about 2 ns. In S45A, this hydrogen bond in chain A broke after 17 ns, but were very stable in other chains. This inconsistency might result from the limited accuracy of the force field utilized. Snapshots were extracted from the 4–17 ns trajectory of S45A for the free energy calculation, while for the other systems the snapshots are taken from 4–20 ns.



**Figure 3.** RMSD of the backbone (BB) and the biotin (Lig) atoms along the simulation for (a) WT, (b) S45A, (c) D128A and (d) DM. Chains are labeled by A to D.



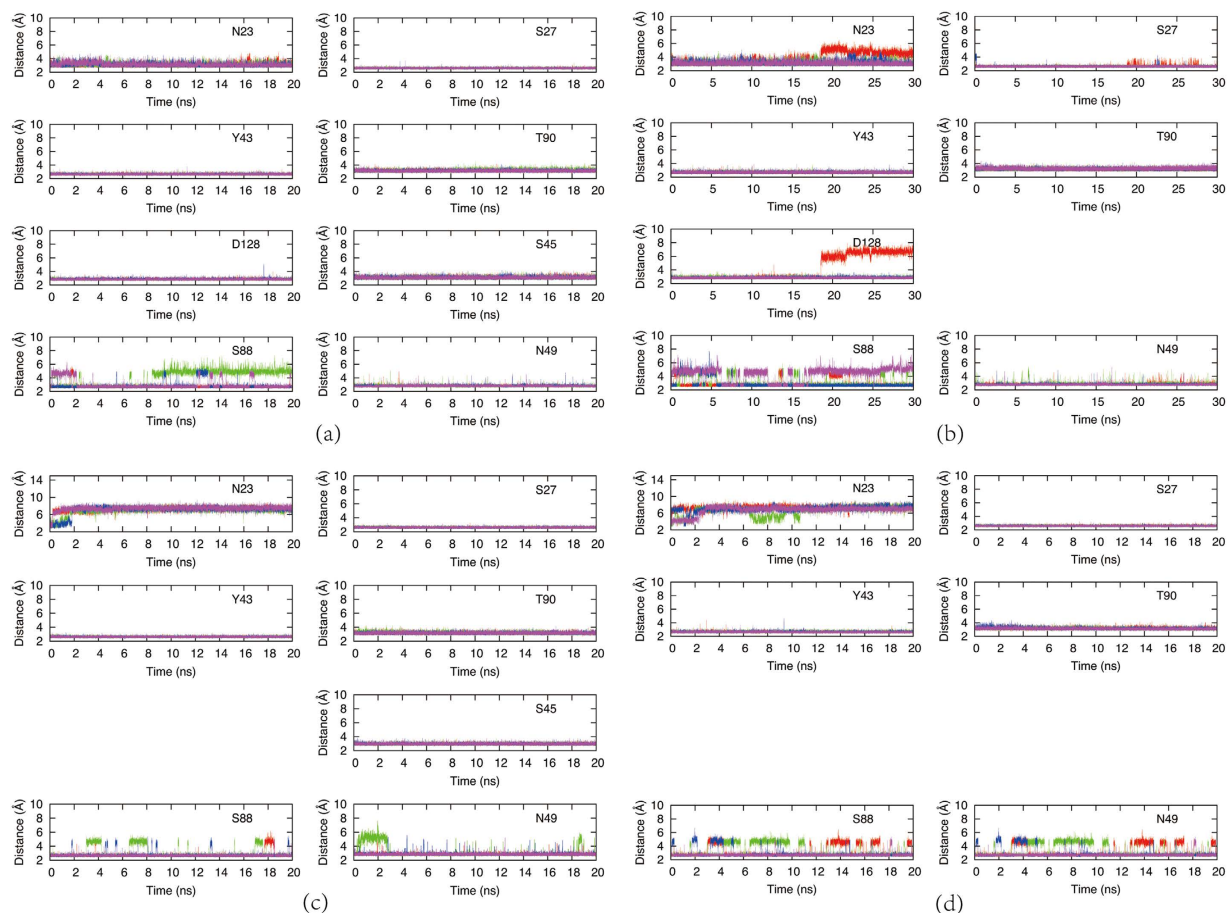
**Figure 4.** The average B-factor values of the residues.



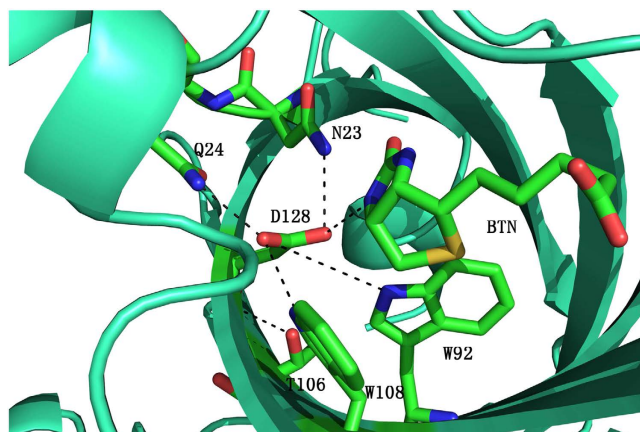
**Figure 5.** The contact maps which show the interaction between the residues of protein. (top right) The difference map between S45A and WT. (bottom left) The difference map between D128A and WT. (bottom right) The difference map between DM and WT.

**MM/PBSA analysis of binding affinity.** The results from single trajectory MM/PBSA analysis are shown in Table 2. Due to the existence of several hydrogen bonds, WT has a strong Coulomb interaction ( $-103.64$  kcal/mol) with biotin, of which about three fourths ( $76.33$  kcal/mol) is canceled by the polar desolvation free energy. The total electrostatic term of enthalpy is  $-27.31$  kcal/mol, which is about the same magnitude as the vdW interaction ( $-30.76$  kcal/mol). The nonpolar solvation free energy contributes to the binding affinity by  $-3.37$  kcal/mol for WT, which is almost invariant in all the four systems. The entropy loss ( $T\Delta S$ ) is also very consistent among all the system, which is in the range from  $18.63$  to  $20.27$  kcal/mol. Without considering the penalty from loop closing, the binding free energy of biotin to WT is  $-41.17$  kcal/mol with a standard error of  $0.36$  kcal/mol. This result is consistent with our previous study<sup>14</sup>. From WT to S45A, a hydrogen bond is eliminated, leading to an  $8.58$  kcal/mol loss of Coulomb interaction between the ligand and the receptor, accompanied by a  $2.74$  kcal/mol drop of polar dehydration free energy. The net loss of binding enthalpy and total binding free energy are  $5.55$  and  $5.27$  kcal/mol, respectively, which are very close to the experimental measurement ( $6.10$  and  $4.30$  kcal/mol). The strong Coulomb repulsion between the carbonyl groups in biotin and D128 is removed after the D128A mutation, which results in an enhancement of Coulomb interaction by  $-25.08$  ( $-128.72$  vs.  $-103.64$ ) kcal/mol. Nonetheless, the desolvation penalty increases by  $32.02$  kcal/mol. The net loss of binding enthalpy and total binding free energy are  $8.49$  and  $6.85$  kcal/mol, respectively, which are about  $2.5$  kcal/mol larger than the experimental values ( $5.90$  and  $4.40$  kcal/mol). From S45A to DM, the Coulomb interaction is strengthened by  $29.33$  kcal/mol, which is larger than that from WT to D128A. In concert, the increase of dehydration penalty is  $37.51$  kcal/mol. The total binding free energy of biotin to DM is  $-26.37$  kcal/mol, which is  $14.80$  kcal/mol weaker than that to WT. This decrease in binding affinity shows an extra loss of  $2.68$  kcal/mol relative to the direct sum of two single mutations, indicating a cooperative characteristics of S45 and D128 in binding to biotin.

Within this extra loss of binding free energy,  $1.54$  kcal/mol comes from the enthalpy and  $1.14$  kcal/mol comes from the entropic contribution (See Table 3). Although the mutations lead to the annihilations of the hydrogen bonds to the ligand, the gas phase interaction, which is defined as the sum of  $\Delta E_{ele}$  and  $\Delta E_{vdw}$ , cannot explain this cooperativity. Indeed, it gives a wrong sign ( $-3.92$  kcal/mol). The enthalpic contribution to this cooperativity



**Figure 6.** The bond length variations of those critical hydrogen bonds which are directly bound to biotin along the simulations for (a) WT, (b) S45A, (c) D128A and (d) DM. N23, S27, Y43, T90, D128, S45, S88, N49 are bound to biotin. Chain A, B, C, D are colored in red, green, blue and purple, respectively. White spaces are left for clarity.



**Figure 7.** The local environment of residue D128 in the WT system. The hydrogen bonds to D128 are marked as dashed lines.

comes from the solvation free energy (5.46 kcal/mol), which is a global property. This result does not agree with the experimental decomposition of the contributions, which shows that the entropy change is the main source of the cooperativity. We believe that the inconsistency originates from the dissimilar definitions of entropy in the experiment and in the MM/PBSA analysis. In the former, the entropy includes the contributions from both of

	$\Delta E_{ele}$	$\Delta E_{vdw}$	$\Delta E_{np}$	$\Delta E_{pb}$	$\Delta E_{es}$	$\Delta H$	$-T\Delta S$	$\Delta G$
WT	-103.64	-30.76	-3.37	76.33	-27.31	-61.44	20.27	-41.17 ± 0.36
S45A	-95.06	-31.02	-3.40	73.59	-21.47	-55.89	19.99	-35.90 ± 0.36
D128A	-128.72	-29.13	-3.45	108.35	-20.36	-52.95	18.63	-34.32 ± 0.37
DM	-124.39	-29.06	-3.51	111.10	-13.29	-45.86	19.49	-26.37 ± 0.37

**Table 2.** Binding free energy components in a unit of kcal/mol. The standard errors of the total binding free energies are also shown.

System	MM/PBSA								ITC		
	$\Delta G$	$\Delta\Delta G$	$\Delta H$	$\Delta\Delta H$	$\Delta\Delta H_{gas}$	$\Delta\Delta G_{sol}$	$-T\Delta S$	$-T\Delta\Delta S$	$\Delta\Delta G$	$\Delta\Delta H$	$-T\Delta\Delta S$
WT	-41.17		-61.44				20.27		0.0	0.0	0.0
S45A	-35.90	5.27	-55.89	5.55	8.32	-2.77	19.99	-0.28	4.3	6.1	-1.8
D128A	-34.32	6.85	-52.95	8.50	-23.45	31.94	18.63	-1.64	4.4	5.9	-1.5
DM (predicted) <sup>1</sup>	-29.05	12.12	-47.39	14.05	-15.13	29.17	18.35	-1.92	8.7	12.0	-3.3
DM (calculated) <sup>2</sup>	-26.37	14.80	-45.86	15.59	-19.05	34.63	19.49	-0.78	10.1	10.3	-0.3
Cooperativity <sup>3</sup>		2.68		1.54	-3.92	5.46		1.14	1.4	-1.6	3.0

**Table 3.** The calculated and experimental variations of binding free energy (kcal/mol) relative to WT. <sup>1</sup>The loss of binding affinity predicted by a direct sum of single-mutant energetic perturbations ( $\Delta E_{WT} + \Delta\Delta E_{S45A} + \Delta\Delta E_{D128A}$ ). <sup>2</sup>Energy calculated by MM/PBSA. <sup>3</sup>(calculated - predicted) for energies.

System	S(A)45	D(A)128	N23	S27	Y43	N49	S88	T90	biotin
WT	-7.46	-14.00	-5.80	-18.31	-18.46	-15.44	-11.73	-5.30	50.27
S45A	-4.17	-13.90	-5.77	-16.22	-17.34	-19.45	-11.81	-3.95	50.12
D128A	-8.23	-0.20	0.27	-16.86	-17.53	-14.32	-14.58	-4.94	46.11
DM	-5.02	-0.15	0.23	-16.92	-17.71	-12.63	-12.23	-5.46	43.47
Cooperativity	-0.08	-0.05	-0.07	-2.16	-1.31	5.70	2.43	-1.87	-2.20

**Table 4.** The decomposition of enthalpy for the eight residues which directly interact with biotin (kcal/mol).

the complex and the solvent molecules. While in MM/PBSA analysis, the solvent contribution to the entropy is merged into the nonpolar solvation term.

**Free energy decomposition.** In order to obtain the fingerprints of the interaction between the ligand and the residues, pairwise energy decompositions of enthalpy were carried out for the eight residues which directly interact with biotin in the WT system<sup>72,73</sup>. The results are listed in Table 4. The interaction between biotin and the 45<sup>th</sup> residue decreases in S45A for losing one hydrogen bond, while the interaction is strengthened in D128A as the biotin moves towards residue S45 once it loses the attraction from residue D128. This can be evidenced by the distance variation between S45 and biotin as shown in Table S1 in Supplementary Information and Fig. 6. The average hydrogen bond length in D128A is 3.0 Å, which is slightly shorter than that in WT (3.1 Å). On the one hand, regardless of whether the 128<sup>th</sup> residue is aspartate or alanine, this S45A mutation leads to a loss of binding affinity between the 45<sup>th</sup> residue and biotin of about 3.2 kcal/mol. As a result, the cooperativity comes from residue S(A)45 is merely -0.08 kcal/mol. On the other hand, the attraction between D(A)128 and biotin is immune to the mutation of S45A. D128 has a strong (anion-dipole) interaction with biotin, and bridges the first and the second binding shells by forming a hydrogen bond network with several residues around. The structure near D128 is very rigid and the S45A mutation has no significant impact on the interaction pattern between D128 and biotin. In WT and S45A, the contributions to the binding affinity from D128 are -14.00 and -13.90 kcal/mol, respectively. After the D128A mutation, A128 has only marginal contribution to the binding affinity (-0.20 kcal/mol in D128A and -0.15 kcal/mol in DM). The variation of interaction between N23 and biotin is correlated with that between D128 and biotin as the hydrogen bond between N23 and biotin disappears once D128 mutates into alanine. This is in accordance with the distances of hydrogen bonds, which are 3.1, 3.2, 7.4, and 7.1 Å in WT, S45A, D128A, and DM respectively (see Table S1). The cooperativity originating from D(A)128 and N23 are -0.05 kcal/mol and -0.07 kcal/mol, respectively. S27 forms a hydrogen bond with the ureido group in biotin, and its contributions to the binding affinity in all the mutants are similar and are 1.5 to 2.0 kcal/mol smaller than that in WT. It contributes -2.16 kcal/mol to the cooperativity. Y43 shows a similar characteristic and contributes -1.30 kcal/mol to the binding affinity. The contribution of T90 to the binding affinity decreases in the single mutants as compared to that in WT, but increases a little bit in DM. It contributes to the cooperativity by -1.87 kcal/mol. Residues N49 and S88 have the largest contribution to the cooperativity (5.70 kcal/mol and 2.43 kcal/mol respectively). From D128A to DM, the S45A mutation leads to a loss in binding affinity of 1.69 kcal/mol from N49.



However, from WT to S45A, the S45A mutation leads to an enhanced interaction between N49 and biotin. This enhancement of attraction is caused by the strengthened interaction between the carboxyl group of biotin and the amide group of N49 as the amide group switches towards the carboxyl group in biotin occasionally and forms a hydrogen bond between them. A direct evidence is the distance between the carbon atom in the carboxyl group of biotin and the nitrogen atom in the amide group of N49 showing in Table S2. The distance becomes shorter obviously in S45A than in the other complexes. The contribution from S88 is quite close in WT, S45A and DM, but is strengthened in D128A. This is in accordance with the hydrogen bond length listed in Table S1. N49 and S88 form hydrogen bonds with the carboxyl group of biotin and are susceptible to the depth of biotin buried in the pocket. All these changes mainly come from the pairwise electrostatic interaction between the ligand and the residues (see Table S3 in Supplementary Information). The biotin molecule itself makes a cooperativity contribution of  $-2.20$  kcal/mol, due to the variation of desolvation penalty.

## Conclusion

Hydrogen bondings in the streptavidin-biotin complex show strong cooperativity. Experiments have shown that S45A/D128A double mutation leads to excess loss of binding affinity than the direct sum of the dual single mutations. In this work, molecular dynamics simulations have been carried out to uncover the origin of the cooperativity. The structures of the wild-type streptavidin-biotin complex and the mutants are stable during the simulations. Only marginal alternations of the internal structural relaxation can be detected after these mutations. Free energy calculations show that the change of the gas phase interaction is not the source of the cooperativity, and it mainly originates from the solvation free energy and the entropic contributions. Interaction fingerprints show that mutations cause the change of the interaction not only between the ligand and the mutated residues, but also that between the ligand and the conserved residues. The residue-decomposed free energy picture shows that the enthalpy contribution to the cooperativity mainly comes from N49 and S88, which interact with the carboxyl group of biotin. It is the strengthened interaction in one of the single mutants between biotin and these two residues that results in a positive contribution to the cooperativity. All these results indicate a synergistical feature of the hydrogen bonding in the streptavidin-biotin complexes.

## References

1. Wells, J. A. Additivity of Mutational Effects in Proteins. *Biochemistry* **29**, 8509–8517 (1990).
2. Mark, A. E. & van Gunsteren, W. F. Decomposition of the Free Energy of a System in Terms of Specific Interactions: Implications for Theoretical and Experimental Studies. *J. Mol. Biol.* **240**, 167–176 (1994).
3. Smith, P. E. & van Gunsteren, W. F. When Are Free Energy Components Meaningful? *J. Phys. Chem.* **98**, 13735–13740 (1994).
4. Brady, G. P. & Sharp, K. A. Decomposition of Interaction Free Energies in Proteins and Other Complex Systems. *J. Mol. Biol.* **254**, 77–85 (1995).
5. Horowitz, A. Double-mutant Cycles: A Powerful Tool for Analyzing Protein Structure and Function. *Fold Des.* **1**, R121–R126 (1996).
6. Hanes, M. S. *et al.* Specificity and Cooperativity at  $\beta$ -lactamase Position 104 in TEM-1/BLIP and SHV-1/BLIP Interactions. *Proteins: Struct., Funct., Bioinf.* **79**, 1267–1276 (2011).
7. Green, N. M. Avidin. 1 The Use of [ $^{14}$ C] Biotin for Kinetic Studies and for Assay. *Biochem. J.* **89**, 585–591 (1963).
8. Green, N. M. Thermodynamics of the Binding of Biotin and Some Analogues by Avidin. *Biochem. J.* **101**, 774–780 (1966).
9. Green, N. M. Avidin. *Adv. Protein Chem.* **29**, 85–133 (1975).
10. Bykhovski, A., Zhang, W. D., Jensen, J. & Woolard, D. Analysis of Electronic Structure, Binding, and Vibrations in Biotin-streptavidin Complexes Based on Density Functional Theory and Molecular Mechanics. *J. Phys. Chem. B* **117**, 25–37 (2013).
11. Hyre, D. E. *et al.* Cooperative Hydrogen Bond Interactions in the Streptavidin-biotin System. *Protein Sci.* **15**, 459–467 (2006).
12. Kuntz, I. D., Chen, K., Sharp, K. A. & Kollman, P. A. The Maximal Affinity of Ligands. *Proc. Natl. Acad. Sci. USA* **96**, 9997–10002 (1999).
13. DeChancie, J. & Houk, K. N. The Origins of Femtomolar Protein-ligand Binding: Hydrogen-bond Cooperativity and Desolvation Energetics in the Biotin-(strept)avidin Binding Site. *J. Am. Chem. Soc.* **129**, 5419–5429 (2007).
14. Zeng, J., Jia, X. Y., Zhang, J. Z. H. & Mei, Y. The F130L Mutation in Streptavidin Reduces Its Binding Affinity to Biotin Through Electronic Polarization Effect. *J. Comput. Chem.* **34**, 2677–2686 (2013).
15. Stenkamp, R. E., Le Trong, I., Klumb, L., Stayton, P. S. & Freitag, S. Structural Studies of the Streptavidin Binding Loop. *Protein Sci.* **6**, 1157–1166 (1997).
16. Chu, V., Stayton, P. S., Freitag, S., Le Trong, I. & Stenkamp, R. E. Thermodynamic and Structural Consequences of Flexible Loop Deletion by Circular Permutation in the Streptavidin-biotin System. *Protein Sci.* **7**, 848–859 (1998).
17. Chelvaraja, S. & Meirovitch, H. Calculation of the Entropy and Free Energy by the Hypothetical Scanning Monte Carlo Method: Application to Peptides. *J. Chem. Phys.* **122**, 054903 (2005).
18. Chelvaraja, S. & Meirovitch, H. Calculation of the Entropy and Free Energy of Peptides by Molecular Dynamics Simulations Using the Hypothetical Scanning Molecular Dynamics Method. *J. Chem. Phys.* **125**, 024905 (2006).
19. Freitag, S. *et al.* Structural Studies of Binding Site Tryptophan Mutants in the High-affinity Streptavidin-biotin Complex. *J. Mol. Biol.* **279**, 211–221 (1998).
20. Klumb, L. A., Chu, V. & Stayton, P. S. Energetic Roles of Hydrogen Bonds at the Ureido Oxygen Binding Pocket in the Streptavidin-biotin Complex. *Biochemistry* **37**, 7657–7663 (1998).
21. Stayton, P. S. *et al.* Streptavidin-biotin Binding Energetics. *Biomol. Eng.* **16**, 39–44 (1999).
22. Hyre, D. E., Le Trong, I., Freitag, S., Stenkamp, R. E. & Stayton, P. S. Ser45 Plays an Important Role in Managing Both the Equilibrium and Transition State Energetics of the Streptavidin-biotin System. *Protein Sci.* **9**, 878–885 (2000).
23. Freitag, S. *et al.* A Structural Snapshot of an Intermediate on the Streptavidin-biotin Dissociation Pathway. *Proc. Natl. Acad. Sci. USA* **96**, 8384–8389 (1999).
24. Jelasarov, I. & Bosshard, H. R. Isothermal Titration Calorimetry and Differential Scanning Calorimetry as Complementary Tools to Investigate the Energetics of Biomolecular Recognition. *J. Mol. Recognit.* **12**, 3–18 (1999).
25. Lifson, S. & Warshel, A. Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and N-alkane Molecules. *J. Chem. Phys.* **49**, 5116–5129 (1968).
26. Brooks, B. R. *et al.* CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **4**, 187–217 (1983).
27. Pearlman, D. A. *et al.* AMBER, A Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **91**, 1–41 (1995).

28. Hagler, A. T., Huler, E. & Lifson, S. Energy Functions for Peptides and Proteins. I. Derivation of a Consistent Force Field Including the Hydrogen Bond from Amide Crystals. *J. Am. Chem. Soc.* **96**, 5319–5327 (1974).
29. Nemethy, G., Pottle, M. S. & Scheraga, H. A. Energy Parameters in Polypeptides. 9. Updating of Geometrical Parameters, Nonbonded Interactions, and Hydrogen Bond Interactions for the Naturally Occurring Amino Acids. *J. Phys. Chem.* **87**, 1883–1887 (1983).
30. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
31. Weiner, S. J. *et al.* A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).
32. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **7**, 230–252 (1986).
33. Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).
34. Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **17**, 490–519 (1996).
35. Warshel, A. Calculations of Enzymic Reactions: Calculations of  $pK_a$ , Proton Transfer Reactions, and General Acid Catalysis Reactions in Enzymes. *Biochemistry* **20**, 3167–3177 (1981).
36. Schaefer, S., Van Vlijmen, H. W. T. & Karplus, M. Electrostatic Contributions to Molecular Free Energies in Solution. In *Linkage Thermodynamics of Macromolecular Interactions* vol. 51 of *Adv. Protein Chem.* 1–57 (Academic Press, 1998).
37. Simonson, T. Electrostatics and Dynamics of Proteins. *Rep. Prog. Phys.* **66**, 737 (2003).
38. Ji, C. G., Mei, Y. & Zhang, J. Z. H. Developing Polarized Protein-specific Charges for Protein Dynamics: MD Free Energy Calculation of  $pK_a$  Shifts for Asp26/Asp20 in Thioredoxin. *Biophys. J.* **95**, 1080–1088 (2008).
39. Zeng, J., Duan, L. L., Zhang, J. Z. H. & Mei, Y. A Numerically Stable Restrained Electrostatic Potential Charge Fitting Method. *J. Comput. Chem.* **34**, 847–853 (2013).
40. Ji, C. G. & Mei, Y. Some Practical Approaches to Treating Electrostatic Polarization of Proteins. *Accounts Chem. Res.* **47**, 2795–2803 (2014).
41. Tong, Y., Mei, Y., Li, Y. L., Ji, C. G. & Zhang, J. Z. H. Electrostatic Polarization Makes a Substantial Contribution to the Free Energy of Avidin-biotin Binding. *J. Am. Chem. Soc.* **132**, 5137–5142 (2010).
42. Mei, Y., Li, Y. L., Zeng, J. & Zhang, J. Z. H. Electrostatic Polarization is Critical for the Strong Binding in Streptavidin-biotin System. *J. Comput. Chem.* **33**, 1374–1382 (2012).
43. Jia, X. Y., Zeng, J., Zhang, J. Z. H. & Mei, Y. Accessing the Applicability of Polarized Protein-specific Charge in Linear Interaction Energy Analysis. *J. Comput. Chem.* **35**, 737–747 (2014).
44. Gilson, M. K. & Zhou, H. X. Calculation of Protein-ligand Binding Affinities. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 21–42 (2007).
45. Zwanzig, R. W. High-temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **22**, 1420–1426 (1954).
46. Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **3**, 300–313 (1935).
47. Straatsma, T. P. & McCammon, J. A. Multiconfiguration Thermodynamic Integration. *J. Chem. Phys.* **95**, 1175–1188 (1991).
48. Kästner, J., Senn, H. M., Thiel, S., Otte, N. & Thiel, W. QM/MM Free-energy Perturbation Compared to Thermodynamic Integration and Umbrella Sampling: Application to an Enzymatic Reaction. *J. Chem. Theory Comput.* **2**, 452–461 (2006).
49. Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Chem. Phys.* **22**, 245–268 (1976).
50. Shirts, M. R. & Pande, V. S. Comparison of Efficiency and Bias of Free Energies Computed by Exponential Averaging, the Bennett Acceptance Ratio, and Thermodynamic Integration. *J. Chem. Phys.* **122**, 144107 (2005).
51. Åqvist, J., Medina, C. & Samuelsson, J. E. A New Method for Predicting Binding Affinity in Computer-aided Drug Design. *Protein Eng.* **7**, 385–391 (1994).
52. Genheden, S. & Ryde, U. Comparison of the Efficiency of the LIE and MM/GBSA Methods to Calculate Ligand-binding Energies. *J. Chem. Theory Comput.* **7**, 3768–3778 (2011).
53. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A. & Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. *J. Am. Chem. Soc.* **120**, 9401–9409 (1998).
54. Brown, S. P. & Muchmore, S. W. High-throughput Calculation of Protein-ligand Binding Affinities: Modification and Adaptation of the MM-PBSA Protocol to Enterprise Grid Computing. *J. Chem. Inf. Model.* **46**, 999–1005 (2006).
55. Gouda, H., Kuntz, I. D., Case, D. A. & Kollman, P. A. Free Energy Calculations for Theophylline Binding to an RNA Aptamer: Comparison of MM-PBSA and Thermodynamic Integration Methods. *Biopolymers* **68**, 16–34 (2003).
56. Miyamoto, S. & Kollman, P. A. Absolute and Relative Binding Free-Energy Calculations of the Interaction of Biotin and Its Analogs with Streptavidin Using Molecular-Dynamics Free-Energy Perturbation Approaches. *Proteins: Struct. Funct. Bioinf.* **16**, 226–245 (1993).
57. Wang, J., Dixon, R. & Kollman, P. A. Ranking ligand binding affinities with avidin: A molecular dynamics-based interaction energy study. *Proteins* **34**, 69–81 (1999).
58. Kuhn, B. & Kollman, P. A. Binding of a diverse set of ligands to avidin and streptavidin: An accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *J. Med. Chem.* **43**, 3786–3791 (2000).
59. Brown, S. P. & Muchmore, S. W. High-throughput calculation of protein-ligand binding affinities: Modification and adaptation of the MM-PBSA protocol to enterprise grid computing. *J. Chem. Inf. Model.* **46**, 999–1005 (2006).
60. Kuhn, B. & Gerber, P. Validation and use of the MM-PBSA approach for drug discovery. *J. Med. Chem.* **48**, 4040–4048 (2005).
61. Weis, A., Katebzadeh, K., Soderhjelm, P., Nilsson, I. & Ryde, U. Ligand affinities predicted with the MM/PBSA method: Dependence on the simulation method and the force field. *J. Med. Chem.* **49**, 6596–6606 (2006).
62. Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **51**, 69–82 (2011).
63. Fiser, A., Do, R. K. G. & Sali, A. Modeling of Loops in Protein Structures. *Protein Sci.* **9**, 1753–1773 (2000).
64. Fiser, A. & Sali, A. ModLoop: Automated Modeling of Loops in Protein Structures. *Bioinformatics* **19**, 2500–2501 (2003).
65. Duan, Y. *et al.* A Point-charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-phase Quantum Mechanical Calculations. *J. Comput. Chem.* **24**, 1999–2012 (2003).
66. Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
67. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **79**, 926–935 (1983).
68. Gaussian 09, Revision B.01 (2010). Gaussian, Inc.: Wallingford, Connecticut, United States. URL <http://www.gaussian.com/>.
69. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
70. Darden, T., York, D. M. & Pedersen, L. G. Particle Mesh Ewald: An  $N \log(N)$  Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
71. Sridharan, S., Nicholls, A. & Sharp, K. A. A Rapid Method for Calculating Derivatives of Solvent Accessible Surface-areas of Molecules. *J. Comput. Chem.* **16**, 1038–1044 (1995).

72. Gohlke, H., Kiel, C. & Case, D. A. Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *J. Mol. Biol.* **330**, 891–913 (2003).
73. Metz, A. *et al.* Hot Spots and Transient Pockets: Predicting the Determinants of Small-Molecule Binding to a Protein-Protein Interface. *J. Chem. Inf. Model.* **52**, 120–133 (2012).

### Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant Nos 21173082 and 21433004) and the Large Instruments Open Foundation of East China Normal University. We thank Supercomputer Center of East China Normal University for CPU time support.

### Author Contributions

Y.M. conceived and designed the experiment. F.L. performed the experiment. F.L. and Y.M. analyzed the data. F.L., J.Z.H.Z. and Y.M. wrote the paper. All authors reviewed the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Liu, F. *et al.* The origin of the cooperativity in the streptavidin-biotin system: A computational investigation through molecular dynamics simulations. *Sci. Rep.* **6**, 27190; doi: 10.1038/srep27190 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>