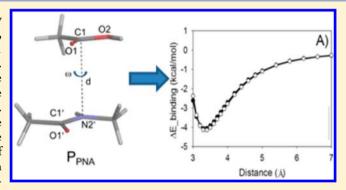
Carboxyl—Peptide Plane Stacking Is Important for Stabilization of Buried E305 of Trichoderma reesei Cel5A

Chunyan He, Jingfei Chen, Liaoyuan An, Yefei Wang, Zhiyu Shu, and Lishan Yao*

Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266061, China

Supporting Information

ABSTRACT: Hydrogen bonds or salt bridges are usually formed to stabilize the buried ionizable residues. However, such interactions do not exist for two buried residues D271 and E305 of Trichoderma reesei Cel5A, an endoglucanase. Mutating D271 to alanine or leucine improves the enzyme thermostability quantified by the temperature T_{50} due to the elimination of the desolvation penalty of the aspartic acid. However, the same mutations for E305 decrease the enzyme thermostability. Free energy calculations based on the molecular dynamics simulation predict the thermostability of D271A, D271L, and E305A (compared to WT) in line with the experimental observation but overestimate the thermostability of E305L. Quantum mechanical calculations suggest



that the carboxyl-peptide plane stacking interactions occurring to E305 but not D271 are important for the carboxyl group stabilization. For the protonated carboxyl group, the interaction energy can be as much as about -4 kcal/mol for parallel stacking and about -7 kcal/mol for T-shaped stacking. For the deprotonated carboxyl group, the largest interaction energies for parallel stacking and T-shaped stacking are comparable, about -7 kcal/mol. The solvation effect generally weakens the interaction, especially for the charged system. A search of the carboxyl-peptide plane stacking in the PDB databank indicates that parallel stacking but not T-shaped stacking is quite common, and the most probable distance between the two stacking fragments is close to the value predicted by the QM calculations. This work highlights the potential role of carboxyl amide $\pi-\pi$ stacking in the stabilization of aspartic acid and glutamic acid in proteins.

■ INTRODUCTION

Ionizable residues inside proteins are important for functions, such as ligand binding and enzymatic catalysis. However, the desolvation penalty of transferring an ionizable group from water to the protein interior can greatly destabilize the residue and the protein native state. As a result, the ionizable residues tend to be neutral in the protein interior to reduce the desolvation energy, so that the pK_a values of acidic (basic) groups are usually higher (lower) than the values of exposed ionizable residues. Even so, burying an ionizable residue is generally considered as destabilizing.^{1–3} Statistical analysis of protein structures in the PDB databank suggests that buried ionizable residues generally form hydrogen bonds or salt bridges with surrounding residues,4 and the buried ionizable residues are more common in larger proteins.⁵ The polar and charge interactions provide a preorganized environment⁶ important for stabilizing the ionizable residues in various protonation states.3,7,8

Tr. Cel5A is a key component of endoglucanases produced by T. reesei. The X-ray structure of the Cel5A catalytic domain (CD) adopts a $(\alpha/\beta)_8$ TIM-barrel fold. Two ionizable residues, D271 and E305, are buried but surprisingly do not form hydrogen bonds or salt bridges with the surrounding protein atoms (Figure S1, Supporting Information). In this work, we show that mutating D271 to a hydrophobic residue, for example, alanine or leucine, improves the thermostability of the enzyme, whereas mutating E305 to the same type of hydrophobic residues destabilizes the enzyme. Further inspection suggests that unlike D271, the carboxyl of E305 forms parallel stacking with the backbone peptide plane made by W292-G293 and T-shaped stacking with the peptide plane of G291-W292. Quantum mechanical (QM) calculations of a model system demonstrate that the stacking interactions between the carboxyl group and the peptide plane can be as strong as -4 to -7 kcal/mol, depending on the relative orientation of the two groups and their distance as well as the protonation state of the carboxyl group. The negatively charged carboxyl group tends to give a stronger interaction than the neutral one for parallel stacking. The stacking, essentially the carboxyl amide π – π interaction, is different from the hydrogen bonding or the salt bridge, which is the known stabilizing interaction for ionizable side chains. A survey of the PDB databank shows that carboxyl amide $\pi - \pi$ parallel stacking

Received: October 9, 2014 Published: January 8, 2015



104

occurs in various proteins, underlining the importance of stacking in the stabilization of glutamate and aspartate.

■ METHODS AND MATERIALS

Molecular Dynamics Simulations. Molecular dynamics (MD) simulations were carried out using Gromacs 4.5. 11,12 with the Amber ff99SB force field¹³ and TIP3P water. ¹⁴ The starting Cel5A structures were from X-ray crystallography (WT, pdb: 3QR3¹⁰). The mutant structure was created using the software Rosetta3, 15 where only the side chain of the mutated residue was optimized and all other atoms were fixed to the Xray structure of WT. The pK, values of ionizable residues were predicted by using the program PROPKA3.16 The residues with pK_a values smaller than (larger than or equal to) 5.0 were assigned as deprotonated (protonated). The protein was solvated by adding 12.0 Å TIP3P water in a rectangular box, and counterions were used to neutralize the system using the Leap program in the Amber 11 software.¹⁷ Before free energy calculations, 1000 steps of energy minimization followed by 1 ns MD simulation at constant pressure (1 atm) and temperature (300 K) were performed to equilibrate the system. The box dimensions were 77 Å \times 69 Å \times 77 Å after equilibration. The pressure was regulated using the extended ensemble Parrinello-Rahman approach, 18,19 and the temperature was controlled by a modified Berendsen thermostat.²⁰ The particle mesh Ewald method^{21,22} was used to evaluate the contributions of the long-range electrostatic interactions. A nonbonded pair list cutoff of 10.0 Å was used, and the nonbonded pair list was updated every 0.01 ps. All bonds to hydrogen atoms in the protein were constrained by using the LINCS²³ algorithm, whereas bonds and angles of water molecules were constrained by the SETTLE²⁴ algorithm, allowing a time step of 0.002 ps. Fifty MD snapshots from the last 500 ps simulation were analyzed. The hydrogen bond was assigned when the distance of the two heavy atoms is <3.5 Å and the angle (hydrogen-donor-acceptor) is <30°.

The p K_a calculations of D271 and E305 generally followed the procedure by Case and co-workers.²⁵ The Amber ff99SB charges were modified for the aspartic acid or the glutamic acid to create the protonated state and the ionized state (Figure S2, Supporting Information). The bonding parameters were not changed. Because the same procedure is used for both the protein and the reference state, the free energy contribution from the bonding parameters should largely cancel out. For the free energy simulation of the charge transformation (for D271 and E305) in Cel5A, 32 λ -windows were built ($\lambda = 0.05n$ (if λ ≤ 0.8), $\lambda = 0.8 + 0.02n$ (if $0.8 < \lambda \leq 0.9$), and $\lambda = 0.9 + 0.01n$ (if $0.9 < \lambda \le 1$), where *n* is an integer) with $\lambda = 0$ (1) corresponding to the protonated (deprotonated) state. Also, 300 ps Hamiltonian replica exchange MD (REMD) simulations²⁶ were run for the 32 λ -windows simultaneously, with the exchange attempted once every 1000 steps. The last 200 ps data were extracted for the free energy evaluation using the Bennett's acceptance ratio method.²⁷ The 200 ps simulation data were divided to five even blocks to calculate the standard deviation of corresponding ΔG ($\Delta G = G$ (deprotonated) – G(protonated)) values, which is used as the estimation for the error of ΔG . In the MD simulation of WT Cel5A, the carboxyl side chains of D271 and E305 did not flip so that the two oxygen atoms are not equivalent due to the local environmental differences. To account for the differences, the protonated state of D271 (E305) was created by adding a proton to either OD1 (OE1) or OD2 (OE2) (Figure S2, Supporting Information).

For the reference state, a small tripeptide GXG (solvated in water with the box dimensions same as those of the protein system) was used where X is the ionizable residue. A similar simulation procedure was adopted for GXG except that the conventional MD simulation was performed for each λ –window system. The double free energy difference $\Delta\Delta G = \Delta G(\text{protein}) - \Delta G(\text{GXG})$ is related to the p K_a of the ionizable group by the equation p $K_a(\text{protein}) = pK_a(\text{GXG}) + 0.434\Delta\Delta G/kT$, where k is the Boltzmann constant and T is the temperature.

The folding free energy difference between a mutant X and Y (another mutant or WT) was calculated from the difference of the free energies between the folded and unfolded simulations (Figure 1). The details of the calculations were described

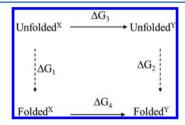


Figure 1. Thermodynamic cycle for folding free energy calculations. ΔG_1 (ΔG_2) is the folding free energy of the mutant X (Y). The folding free energy difference $\Delta \Delta G = \Delta G_2 - \Delta G_1$ can also be written as $\Delta \Delta G = \Delta G_4 - \Delta G_3$, where ΔG_3 and ΔG_4 were evaluated using MD free energy simulations.

previously.²⁸ Briefly, a thermodynamic circle was built where the folding free energy difference was derived from the alchemical transformation of the mutant X to Y^{29-32} The transformation from X to Y was separated into two steps, $X \rightarrow I$ and $Y \rightarrow I$, where I is an intermediate state corresponding to an alanine at the mutated site, and the free energy difference of the two steps is the $X \to Y$ transformation free energy. Taking $X \to Y$ I alchemy as an example, a λ -dependent Hamiltonian H(λ) was introduced to remove the side chain charges and then annihilate the redundant side chain atoms. Each process, consisting of 32 λ -windows (with the λ parameters the same as in the pK_a simulation), was run for 300 ps with the REMD method, and the last 200 ps data were used for the free energy evaluation.²⁷ To ensure a proper convergence, the $dH/d\lambda$ values at different λ -windows were examined. Figure S3 of the Supporting Information shows some typical plots of $dH/d\lambda$ values versus time, and they appeared stable during the REMD simulation. The exchange rate in the REMD simulations was typically about 70% or higher, which allowed the system to reach the equilibrium quickly. The unfolded state was modeled with a GXG tripeptide, which has been shown to yield results in a good agreement with the experimental $\Delta \Delta G$ values.³³ Similar to the folded state, the two step transformations were used to calculate the free energy difference for the unfolded state.

Quantum Mechanical Calculations. Model compounds, formic acid (protonated), acetic acid (both protonated and deprotonated), and N-methylacetamide (NMA), were optimized at the MP2/6-31+G**^{34,35} level, and single point energies were calculated at the MP2/aug-cc-pvdz level. The two molecules (formic acid and NMA or acetic acid and NMA) were stacked in a parallel or T-shaped configuration. The energy of the complex was calculated at MP2/aug-cc-pvdz. The acetic acid and NMA binding energy is defined by $\Delta E_{\rm bind} = E_{\rm complex} - (E_{\rm NMA} + E_{\rm acet}) + \Delta E_{\rm BSSE}$, where $\Delta E_{\rm BSSE}$ is the

counterpoise correction 38 to eliminate the basis set superposition error ($\Delta E_{\rm BSSE} = E_{\rm NMA} - E_{\rm NMA}^{} + E_{\rm acet}^{} - E_{\rm acet}^{} *$, asterisk (*) representing monomers calculated with "ghost" orbitals). A similar expression was used to calculate the formic acid NMA binding energy. The relative orientation of the two molecules and their distance were varied to study the geometric effect on the binding energy. All the QM calculations were performed by Gaussian 09. 39

Cloning, Expression, and Purification of Tr. Cel5A. The DNA encoding residues of the Cel5A catalytic domain (Cel5A-CD) from T. reesei QM9414 and a 9*His tag at the C-terminus was ligated with the vector pET-22b. The ligation mixture was transformed into an E. coli strain DH10B. The expression vector (pET-22b-Cel5A-CD) was then transformed into the E. coli strain Rosetta-gami (DE3). All mutations were made by PCR-based site-directed mutagenesis and verified by DNA sequencing. All the mutants were expressed and purified in a similar way. Briefly, 250 mL of LB medium containing 100 μ g/ mL Ampicillin was inoculated with a fresh colony of expression strain Rosetta-gami (DE3) containing pET-22b-Cel5A-CD. The culture was grown at 37 °C. When the OD600 of the culture reached 0.8-1.2, a final concentration of 1 mM of IPTG was added to induce the expression of the protein at 16 °C and for 24 h. The cells were harvested by centrifugation, suspended, and lysed by ultrasonication. The lysed cells were centrifuged, and the resulting supernatants were purified by Ni-NTA affinity chromatography (Novagen). The purity was determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The protein concentration was determined by measuring UV absorption at 280 nm, with an extinction coefficient of 67880 M⁻¹ cm⁻¹, calculated from the amino acid composition by using the online tool ProtParam (http://web. expasy.org/protparam/).

 T_{50} Measurements. T_{50} was defined as the temperature at which a 5 min incubation causes a loss of 50% of the activity relative to a reference sample that does not undergo incubation. Twenty-seven microliter samples containing 0.056 μM Tr. Cel5A-CD each were incubated at different temperatures for 5 min. The residual activity against carboxymethyl cellulose (CMC) was measured at 50 °C. The reaction system had a total volume 30 μL containing 0.1% (w/v) CMC and 0.05 μM enzyme (final concentrations) in a 50 mM NaAc, 50 mM NaCl buffer (pH 5.0), and the reaction was stopped after 5 min. The reducing sugar was measured by the PAHBAH assay. The residual activity versus temperature was plotted and fitted using a four-parameter sigmoidal curve.

PDB Databank Searches. We searched the PDB databank (in July 2014) for the high quality protein structures using the program PISCES. The R-factor of each structure was limited to a maximum value of 0.3 (resolution <2.0 Å). The sequence identity between pair structures was limited to <25%. The minimum chain length is 100 amino acids. A total of 6338 protein structures was selected. The carboxyl (of glutamic acid and aspartic acid) amide peptide plane stacking from the selected structures was identified using an in-house VMD⁴³ script.

■ RESULTS AND DISCUSSION

 pK_a Prediction of D271 and E305. There are two buried carboxyl groups D271 and E305 that do not form hydrogen bonds with other protein atoms in Tr. Cel5A (Figure S1, Supporting Information). As predicted by the program POPS, 44 D271 has a solvent accessibility of 14%, which is a fraction of

burial defined by the ratio of solvent accessible area of a residue in a protein to the fully exposed solvent accessible area of this type of residue. E305 has a solvent accessibility of 11%. The pK_a prediction by MD free energy calculations yields a $\Delta\Delta G$ of about 10 kcal/mol for D271 and about 40 kcal/mol for E305, suggesting that both residues are protonated (Table 1). The

Table 1. $\Delta\Delta G$ Predictions of D271 and E305 Protonation Using the MD Free Energy Method

		GXG (kcal/mol)	Cel5A (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
D271	OD1 ^a	-73.8 ± 0.2	-63.4 ± 0.2	10.4 ± 0.3
	$\mathrm{OD2}^a$	-73.8 ± 0.2	-63.4 ± 0.1	10.4 ± 0.2
E305	$OE1^a$	-75.8 ± 0.2	-37.8 ± 2.5	38.0 ± 2.5
	$OE2^a$	-75.8 ± 0.2	-34.5 ± 0.7	41.3 ± 0.7

^aThe atom to which the proton is attached (Figure S1, Supporting Information).

much larger $\Delta\Delta G$ for E305 is consistent with its lower solvent accessibility and the more hydrophobic environment. Two protonation states, corresponding to the proton added either to OD1 (OE1) or OD2 (OE2) of D271 (E305) (Figure S1, Supporting Information), give slightly different $\Delta\Delta G$. For E305, the protonation of O_{e2} is 3.3 kcal/mol more stable than the protonation of $O_{\varepsilon 1}$ (Table 1). The corresponding population of $O_{\varepsilon 2}H$ is greater than 99%. Thus, the $O_{\varepsilon 2}H$ state was selected in the subsequent folding free energy calculations. For D271, the folding free energy calculations were calculated for both states and averaged because the two states have the same $\Delta\Delta G$. The MD trajectory with protonated D271 (or E305) provides more information about the carboxyl interactions with the surroundings. In the MD simulation with the protonated D271 ($O_{\delta 2}$ H as an example), hydrogen bonds are formed between the D271 side chain carboxyl and water molecules. D271 acts as a donor with a hydrogen bond percentage of 100% and an acceptor with 110%. The percentage is defined by the total number of hydrogen bonds formed with water in the MD snapshots divided by the number of snapshots. Because the carboxyl can form multiple hydrogen bonds with different water molecules, the percentage can be larger than 100%. An E305 side chain carboxyl also forms hydrogen bonds (as a donor) with water molecules (88%). The energy decomposition analysis shows that the electrostatic interaction energy between D271 (E305) and its surroundings is $-36.1 \pm 3.6 (-49.7 \pm 4.2)$ kcal/mol, and the corresponding van der Waals energy is $-22.2 \pm 2.2 (-19.5 \pm 2.5)$ kcal/mol.

Mutational Effect on Protein Thermostability. The folding free energy difference between mutant X and Y $\Delta \Delta G$ is defined as $\Delta G_2 - \Delta G_1$, where ΔG_1 (ΔG_2) is the folding free energy of X (Y) (Figure 1). The direct folding free energy calculation is difficult. Instead, a thermodynamic cycle is built so that $\Delta \Delta G$ can be written as $\Delta \Delta G = \Delta G_4 - \Delta G_3$, where ΔG_3 and ΔG_4 are the transformation free energies (from X to Y) of the unfolded state and the folded state, respectively (Figure 1; Table S1, Supporting Information). ΔG_3 (ΔG_4) is evaluated by the MD chemical alchemy simulation, which in this work is composed of two steps: the electrostatic transformation followed by the van der Waals (VDW) transformation. The total free energy difference is $\Delta \Delta G = \Delta \Delta G_{\text{ele}} + \Delta \Delta G_{\text{VDW}}$. The MD free energy calculation of D271A mutation shows that the electrostatic term $\Delta \Delta G_{\text{ele}}$ is negative, but the VDW term $\Delta\Delta G_{\text{VDW}}$ is positive (Table 2). The negative electrostatic $\Delta \Delta G$ is essentially due to the solvation gain when transferring

Table 2. Predicted Folding Free Energy Changes and Experimental T_{50} Changes Caused by Mutation

	$\Delta\Delta G_{ m elec} \ m (kcal/mol)$	$\Delta\Delta G_{ m VDW} \ m (kcal/mol)$	$\Delta\Delta G_{_ ext{tot}} \ ext{(kcal/mol)}$	ΔT_{50}^{a} (°C)
D271A	-2.3 ± 0.3	2.5 ± 0.3	0.2 ± 0.4	0.3 ± 0.2
D271L	-1.9 ± 0.2	-0.8 ± 0.4	-2.7 ± 0.4	2.8 ± 0.2
E305A	-3.1 ± 0.7	5.9 ± 0.4	2.8 ± 0.9	-8.7 ± 0.1
E305L	-2.4 ± 0.7	0.9 ± 0.8	-1.5 ± 1.1	-1.4 ± 0.1
$^a\Delta T_{50}=$	$T_{50}(\text{mutant})$ –	$T_{50}(WT)$, with	$T_{50}(WT) = 72$	$.2 \pm 0.1$ °C.

the carboxyl group from the protein interior (D271) to the solvent (as in the unfolded state, Figure 1). The positive VDW is due to the loss of the contacts between the A271 side chain (as compared to D271) and its surroundings in the protein. So the mutation D271A has a more favorable electrostatic interaction but a less favorable VDW interaction. The two terms largely cancel out (Table 2), yielding a slightly positive $\Delta\Delta G$ (0.2 ± 0.4 kcal/mol), suggesting that the mutation D271A has an insignificant effect on the stability. When D271 is mutated to leucine, the free energy calculations indicate that the $\Delta\Delta G_{\rm VDW}$ is less positive than the D271A mutation due to the larger size of the leucine side chain, but the $\Delta\Delta G_{\rm ele}$ of the D271L mutation is comparable to that of D271A. As a result, the net $\Delta\Delta G$ is negative (-2.7 \pm 0.4 kcal/mol), meaning that D271L is more stable than WT due to the elimination of the aspartate desolvation penalty.

To validate the computational predictions, the thermostability of Cel5A was measured. Unfortunately, Cel5A is unable to recover its activity after 5 min of incubation at high temperature (e.g., 80 °C) suggesting an irreversible unfolding so that the determination of unfolding free energy ΔG experimentally through folding/unfolding equilibrium measurement is not possible. The thermostability of Cel5A is quantified by T_{50} , the incubation temperature at which 50% of enzyme activity is lost (Figure 2). The increased stability of D271L

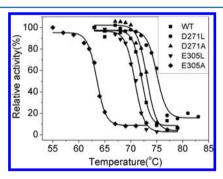


Figure 2. Relative activity at 50 °C after 5 min incubation at different temperatures versus the temperature for WT, D271A, D271L, E305A, and E305L. The data points were fitted to a sigmoid function to estimate the T_{50}^{39} (Table 2).

predicted by MD is confirmed by the experimental measurement (Table 2, Figure 2). As for D271A, experimental T_{50} is larger than WT by 0.3 \pm 0.2 °C, suggesting that this mutant is marginally more stable.

Similar to D271, MD free energy simulations predict that E305A has a negative $\Delta\Delta G_{\rm ele}$ but a positive $\Delta\Delta G_{\rm LVDW}$, and the net $\Delta\Delta G$ is 2.8 \pm 0.9 kcal/mol (Table 2). Due to the loss of the VDW interactions, E305A is a destabilizing mutant, which is consistent with the experimental $T_{\rm 50}$ value where the mutation decreases $T_{\rm 50}$ by 8.7 °C. In comparison, E305L has a

less positive $\Delta\Delta G_{\rm VDW}$ but a $\Delta\Delta G_{\rm ele}$ comparable to E305A so that the predicted $\Delta\Delta G$ is negative (-1.5 \pm 1.1 kcal/mol), suggesting that E305L is more stable than WT. However, the experimental data show that E305L decreases T_{50} by 1.4 °C. Apparently MD simulations underestimate the stability of WT compared to the E305L mutant. Further inspection shows that different from D271 the carboxyl of E305 forms parallel stacking with the peptide plane formed by W292–G293 and T-shaped stacking with the peptide plane formed by G291–W292 (Figure 3). Peptide plane amide—aromatic molecule (e.g.,

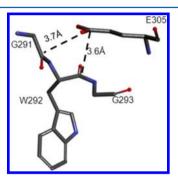


Figure 3. Stacking interactions formed by E305 and G291–W292–G203. The carboxyl of E305 forms parallel stacking with the peptide plane W292–G293 and T-shaped stacking with the peptide plane G291–W292. The correspond distances of 3.6 Å for $C_{\delta}(E305)$ –CO(W292) and 3.7 Å for $C_{\delta}(E305)$ –CO(G291) are from the X-ray structure.

pyridine) stacking has been studied recently, $^{45-47}$ suggesting that amide is a good π -stacking system. The peptide amide—amide stacking interaction is comparable with the intramolecular amide carbonyl hydrogen bonding. However, the carboxyl amide stacking has not been characterized previously. To understand this type of interaction, QM calculations are adopted to study the peptide plane—carboxyl stacking using model complexes, and the results are discussed below.

Carboxyl-Peptide Plane Stacking Interactions. The protonated acetic acid was stacked on top of NMA with the two planes, which are defined by the carboxyl $C_1-O_1-O_2$ of the acetic acid and the NMA $N_2{}^{\prime}-C_1{}^{\prime}-O_1{}^{\prime}$, respectively, aligned in parallell (Figure 4). The proton is attached to the O₂ atom. The vector $C_1-C_1{}'$ was set perpendicular to the peptide plane (model P_{PCA}, short for parallel stacking of protonated carboxyl and NMA peptide plane (with C₁' at the surface center), model A). To quantify the energy dependency on the relative orientation of two molecules, dihedral ω defined as $\angle O_1$ - $C_1 {-} C_1{'} {-} N_2{'}$ was scanned in a 30° increment where all other degrees of freedom were constrained, with the $C_1-C_1{}'$ distance *d* fixed at 3.5 Å. The lowest $\Delta E_{\rm bind}$ of -3.5 kcal/mol, calculated using the MP2/aug-cc-pvdz method, occurs at the ω angle of 0° . Another $\Delta E_{\rm bind}$ minimum of -2.7 kcal/mol can be seen at ω = 150° (Figure 5A). By rotating the acetic acid 180° along the bond C₁-C₂ while fixing NMA, a slightly different parallel stacking model is created (model P_{PCB}). Similar to P_{PCA} , there are two $\Delta E_{\rm bind}$ minima in this model, with a value of -2.3 kcal/ mol at $\omega = -60^{\circ}$ and -4.0 kcal/mol at $\omega = 90^{\circ}$ (Figure 5A). $\Delta E_{\rm bind}$ is positive when ω equals to 30° and 60° for the model P_{PCA}. Further inspection suggests that the positive value results from the steric clash between the methyl group of the acetic acid and the C-methyl group of NMA. To avoid this contact, a slightly different parallel stacking model was built with N2' set at the surface center of the NMA peptide plane and the vector

Figure 4. Stacking models built for the acetic acid and NMA complex: Parallel stacking for protonated (deprotonated) acetic acid, P_{PCA} , P_{PNA} , P_{PCB} , and P_{PCB} and P_{DNA} and P

 $C_1-N_2{}'$ perpendicular to the peptide plane (model P_{PNA}) Figure 4). The dihedral angle ω was scanned with the C₁-N₂ distance d fixed at 3.5 Å, similar as in the model P_{PCA} . ω was redefined by the dihedral angle $\angle O_1 - C_1 - N_2{'} - C_1{'} - 180^\circ$ so that the same ω angle value in two models (P_{PNA} versus P_{PCA}) gives the same relative orientation of the two molecules. The $\Delta E_{\rm bind}$ energy profile is shown in Figure 5A, where the positive values at $\omega = 30^{\circ}$ and 60° disappear. Two $\Delta E_{\rm bind}$ minima of -3.5 and -3.1 kcal/mol are identified at $\omega = 0^{\circ}$ and $\omega = 120^{\circ}$ respectively, similar to the model P_{PCA} . The 180° rotation of the carboxyl group along C₁-C₂ generates another model P_{PNB} (Figure 4). The ΔE_{bind} energy profile of P_{PNB} is similar to that of P_{PCB} with two minima of -3.2 kcal/mol ($\omega = -60^{\circ}$) and -3.6 kcal/mol ($\omega = 90^{\circ}$). The maximum values of ΔE_{bind} are smaller in P_{PNA} and P_{PNB} than those in P_{PCA} and P_{PCB} , suggesting that the steric clashing effect is smaller for the former. To completely eliminate the steric effect caused by the methyl groups, the same parallel stacking models were built for the formic acid and NMA complex. The $\Delta E_{\rm bind}$ energy curves are similar to those in the acetic acid NMA parallel stacking models but with the energy spikes (e.g., in PPCA and PPCB) disappeared (Figure 5B). Meanwhile, the energy minima are also higher (less negative) in the formic acid and NMA models. Apparently, the steric effect caused by the methyl groups alters the energy profile, as expected in proteins where the stacking may render the methylene groups of aspartate or glutamate in contact with the backbone C_{α} (or C_{β}) atoms.

To investigate the distance effect on $\Delta E_{\rm bind}$, the C_1-N_2' distance d was scanned at $\omega=0^\circ$ for the $P_{\rm PNA}$ model. A minimum of -4.1 kcal/mol can be seen at d=3.3 Å (Figure 6A). Favorable stacking interactions persist even at relatively long distances. For example, $\Delta E_{\rm bind}=-1.1$ kcal/mol when d=5.0 Å. A similar $\Delta E_{\rm bind}$ versus C_1-C_1' distance profile was obtained for $P_{\rm PCB}$ with ω fixed at 90° (Figure 6A).

T-shaped acetic acid NMA stacked models were also built with the vector $C_1-C_1{}'$ (model T_{PCA}) or $C_1-N_2{}'$ (model T_{PNA}) perpendicular to the peptide plane (Figure 4). ΔE_{bind} was determined for different ω dihedral angles with the distance d fixed at 4.0 Å ($\omega = \angle O_1 - C_1 - C_1' - N_2'$ and $d = C_1 - C_1'$ distance for T_{PCA} , $\omega = \angle O_1 - C_1 - N_2' - C_1' - 180^\circ$ and $d = C_1 - C_1'$ $C_1 - N_2$ distance for T_{PNA}). A single minimum is observed with $\Delta E_{\rm bind}$ of -5.8 kcal/mol at ω = -60° for $T_{\rm PCA}$ and $\Delta E_{\rm bind}$ of -3.9 kcal/mol at $\omega = -30^{\circ}$ for T_{PNA} (Figure 5C). Thus, it appears that the T_{PCA} model is more stable than T_{PNA} . Inspection of the complex structure indicates that a hydrogen bond is formed between O_1' and O_2H in T_{PCA} (with the O_1' - O_2 distance of 3.5 Å and angle $\angle O_1' - H - O_2$ of 174°). The $\Delta E_{\text{bind}} C_1 - C_1'$ distance (d) scan was performed for T_{PCA} with ω fixed at -60° (Figure 6B). A $\Delta E_{\rm bind}$ minimum of -6.9 kcal/ mol was obtained at the distance d of 3.6 Å, which is 2.8 kcal/ mol lower than the energy minimum for the parallel stacking models P_{PNA} and P_{PCB}. Apparently, the hydrogen bond provides some extra binding energy in T-shaped stacking.

All the models built so far have a protonated carboxyl group, which likely represents the protonation state of most carboxyl groups buried in proteins. However, in some cases, the carboxyl group can be deprotonated, for example, on the protein surface or buried but forming a salt bridge with a lysine or arginine. To investigate the deprotonated carboxyl peptide stacking effect, both parallel and T-shaped stacking models were built for the negatively charged acetic acid and NMA complex (Figure 4). For parallel stacking, two models were built with $C_1{}'$ or $N_2{}'$ of NMA set at the surface center corresponding to $P_{\rm DCA}$ and $P_{\rm DNA}$. The ω dependent $\Delta E_{\rm bind}$ energies were calculated by scanning the dihedral ω where all other degrees of freedom were constrained with the $C_1{}-C_1{}'$ of $P_{\rm DCA}$ (or $C_1{}-N_2{}'$ of $P_{\rm DNA}$) distance d fixed at 3.5 Å. $P_{\rm DCA}$ has two energy minima (Figure 7A), -5.3 kcal/mol ($\omega=-5{}^{\circ}$) and -1.9 kcal/mol ($\omega=145{}^{\circ}$).

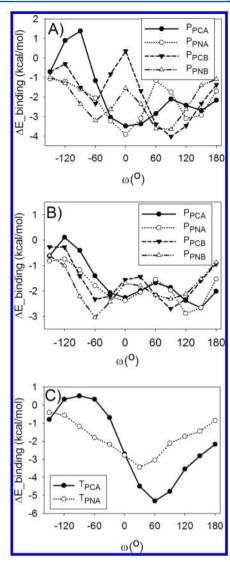


Figure 5. Binding energy $\Delta E_{\rm bind}$ versus dihedral angle ω of protonated acetic acid—NMA parallel stacking models (A), protonated formic acid—NMA parallel stacking models (B), and protonated acetic acid—NMA T-shaped stacking models (C). Distance d was fixed at 3.5 Å.

 P_{DNA} has a ΔE_{bind} energy profile similar to that of P_{DCA} but with lower energy values (Figure 7A), suggesting that for the parallel stacking interaction the C_1 atom of the carboxyl group prefers to align with NMA N_2 ' instead of C_1 '. The two ΔE_{bind} minima

of the model $P_{\rm DNA}$ are -6.6 kcal/mol ($\omega = -35^{\circ}$) and -5.5 kcal/mol ($\omega = 145^{\circ}$). The distance d dependent $\Delta E_{\rm bind}$ profile of $P_{\rm DNA}$ ($\omega = -35^{\circ}$) is shown in Figure 8, with a minimum of -7.0 kcal/mol at d = 3.3 Å.

For T-shaped stacking of the deprotonated carboxyl group, two models were built with C_1-C_1' (model T_{DCA}) or C_1-N_2' (model T_{DNA}) perpendicular to the peptide plane of NMA (Figure 4). The ω -dependent $\Delta E_{\rm bind}$ profiles are shown for both models in Figure 7B, with energies of T_{DNA} about 2 kcal/ mol lower than T_{DCA}, suggesting that the carboxyl group intends to align with N2', similar to that in the parallel stacking model. The two $\Delta E_{\rm bind}$ minima have a similar value of -5.6kcal/mol at ω = -60° and 120° for T_{DNA} , whereas the ΔE_{bind} minima for T_{DCA} are -3.5 kcal/mol at similar ω angles. The distance d-dependent $\Delta E_{\rm bind}$ profile of $T_{\rm DNA}$ was built with ω fixed at 60° , which shows a minimum of -6.5 kcal/mol at d =3.6 Å (Figure 8). ΔE_{bind} of T_{DNA} increases dramatically as the distance is shortened from 3.5 to 3.0 Å, in contrast to the parallel stacking model P_{DNA}, which has a much smaller energy increase.

On the basis of the MP2 calculations, the stacking energy between acetic acid and NMA can be as strong as -4 to -7 kcal/mol, depending on the carboxyl protonation state and the stacking conformation, generally stronger than the amide aromatic ring interaction, 43 which is about -2.5 to -5.5 kcal/ mol. The strong orientational dependence of the amide carboxyl stacking energy for both parallel and T-shaped conformations suggests that the electrostatic interaction (especially for the deprotonated carboxyl group) has a large contribution to the total interaction energy. The importance of electrostatics has been observed in the amide aromatic stacking.⁴³ In principle, the geometry of the monomers can change when they interact, which causes deformation of the structures. This deformation energy, buried in the interaction energy, was not explicitly calculated. However, considering both monomers are rather rigid, the deformation energy is expected to be small. The entropic effect was not considered explicitly in the QM calculations. In a real protein system, this entropic effect may have a nonnegligible contribution to the stacking interaction. The MP2 calculations were performed in the gas phase, and the protein environmental effect on the stacking was neglected. To estimate such an effect, the implicit solvation model PCM⁴⁹ with the solvent chloroform, which has a dielectric constant of 4.7, was adopted to mimic the protein interior effect for the models P_{PNA} and P_{DNA} in the ω angle scan caclulations. It appears that the solvation effect tends to reduce the binding energy $\Delta E_{\rm bind}$ (becomes less negative, Figure S4,

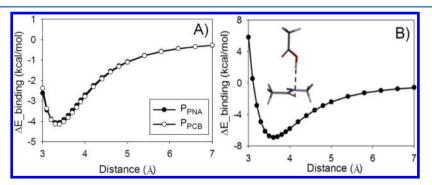


Figure 6. Binding energy $\Delta E_{\rm bind}$ versus distance d of protonated acetic acid-NMA stacking models $P_{\rm PNA}$, $P_{\rm PCB}$ (panel A), and $T_{\rm PCA}$ (panel B). Distance d was scanned with the dihedral angle ω fixed at 0° for $P_{\rm PNA}$, 90° for $P_{\rm PCB}$, and -60° for $T_{\rm PCA}$. The energy minima are -4.1 kcal/mol for $P_{\rm PNA}$ at d of 3.3 Å, -4.1 kcal/mol for $P_{\rm PCB}$ at d of 3.4 Å, and -6.9 kcal/mol for $T_{\rm PCA}$ at d of 3.6 Å.

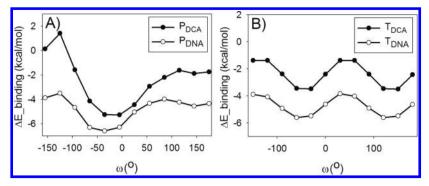


Figure 7. Binding energy $\Delta E_{\rm bind}$ versus dihedral angle ω of deprotonated acetic acid—NMA parallel stacking models (panel A) and deprotonated acetic acid—NMA T-shaped stacking models (panel B). Distance d was fixed at 3.5 Å.

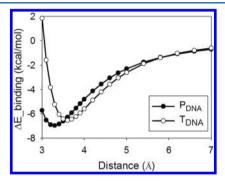


Figure 8. Binding energy $\Delta E_{\rm bind}$ versus distance d of deprotonated acetic acid—NMA stacking models $P_{\rm DNA}$ and $T_{\rm DNA}$. Distance d was scanned with the dihedral angle ω fixed at 0° for $P_{\rm DNA}$, -35° for $P_{\rm DNA}$, and 60° for $T_{\rm DNA}$. Energy minima are -7.0 kcal/mol for $P_{\rm DNA}$ at d of 3.3 Å and -6.5 kcal/mol for $P_{\rm DNA}$ at d of 3.6 Å.

Supporting Information). The $\Delta E_{\rm bind}$ change is relatively small for $P_{\rm PNA}$, about 1 kcal/mol or less, but much larger for $P_{\rm DNA}$, about 4 kcal/mol. The larger solvation effect for $P_{\rm DNA}$ is likely due to the -1 charge of the system.

E305 Side Chain Stacking in the MD Simulation. As discussed above, the X-ray structure shows that the carboxyl group of E305 forms parallel stacking with the W292–G293 peptide plane and T-shaped stacking with the G291–W292 peptide plane (Figure 3). To see whether the stacking interactions persist in the MD simulation, the $C_{\delta}(E305)$ – CO(W292) distance and the $C_{\delta}(E305)$ –CO(G291) distance were calculated in the trajectory of WT Cel5A, with an average value of 3.5 \pm 0.2 Å for the former and 4.1 \pm 0.3 Å for the latter. The $C_{\delta}(E305)$ –CO(G291) distance is about 0.4 Å longer than that from the X-ray structure (Figure 3), suggesting that T-shaped stacking is weakened in the MD simulation.

Although the average $C_{\delta}(E305)$ –CO(W292) distance (3.7 Å) is comparable to the value of the X-ray structure (3.6 Å), the side chain $C_{\beta}-C_{\gamma}-C_{\delta}-O_{\epsilon 1}$ dihedral angle has a fluctuation of about 24°, suggesting that parallel stacking with the W292-G293 peptide plane is not very stable. A weak hydrogen bond is formed (7.7%) between $O_{\epsilon 2}(E305)$ -H and carbonyl O(W292) due to the dihedral rotation in the MD simulation. The molecular mechanics (MM) ff99SB force field¹³ was utilized for the protein in the simulation. To see how well this force field performs in the carboxyl-peptide plane stacking, the binding energies for different protonated acetic acid-NMA parallel stacking models were calculated, including P_{PNA}, P_{PNB}, P_{PCA}, P_{PCB} , T_{PNA} , and T_{PCA} at different ω angles (d fixed at 3.5 Å) and P_{PNA} and P_{PCB} at different distances with the dihedral angle ω fixed at 0° for P_{PNA} and 90° for P_{PCB} (Figure S5, Supporting Information). The overall profiles are similar to those from the QM calculations, in line with a recent study of $\pi - \pi$ stacking between carbon nanotube and amino acid aromatic side chains⁵⁰ but with the binding energies of the optimal configurations about 1-2 kcal/mol more negative than the QM values (Figure S5, Supporting Information). The overestimation of the binding energy by MM calculations, which has been reported for the DNA base pair stacking interactions, is attributed to the anisotropic polarizabilities.⁵¹ It will be interesting to see whether the polarizable force field predicts the binding energies in better agreement with the QM results.

Survey of Carboxyl-Peptide Plane Stacking in Proteins. The carboxyl-peptide plane stacking interactions are pronounced based on the QM calculations of the model system above. To see whether this type of interaction persists generally in protein structures, a search of the backbone peptide plane and aspartic acid (as well as glutamic acid) stacking was performed. Taking aspartic acid as an example, three

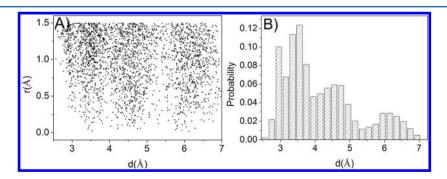


Figure 9. Analysis of carboxyl amide parallel stacking interactions in a database of 6338 protein structures. (A) Scatter plot of the distance *d* versus *r* (see main text for more details). (B) Histogram of distance *d* (probability was corrected for the increase in volume at a larger distance).

parameters are utilized to describe parallel stacking, namely, the angle θ formed between the carboxyl plane (defined by C_v - $O_{\delta 1} - O_{\delta 2}$) and the peptide plane (defined by CO-N-O), the distance d between the center of mass of C_{γ} – $O_{\delta 1}$ – $O_{\delta 2}$ and the center of mass of CO-N-O, and the distance r between the projection of C_v on the peptide plane and the geometric center of atoms CO and N. A similar definition is used for the glutamic acid. The values used to define the parallel stacking interaction are d < 5.0 Å, $\theta < 20^{\circ}$, and r < 1.5 Å. A total of 1793 parallel stacking interactions were identified from a database of 6338 high resolution X-ray structures, yielding an average of 0.28 interactions per structure (Figure 9A). An analysis of the distribution of the distance d shows a maximum occurrence at d = 3.5 Å (Figure 9B), consistent with the QM calculations (Figures 5 and 7). Two more maxima with smaller probabilities also occur at d of 4.7 and 6.1 Å (Figure 9B). Because the QM calculations only have the binding energetic contribution, the agreement on the optimal distance between the statistical analysis and QM calculations suggests that for parallel stacking the energetic contribution dominates the parallel stacking interaction. For T-shaped stacking, the cutoff values are d < 5.0Å, $\theta > 70^{\circ}$, and r < 1.5 Å. Compared to parallel stacking, the occurrence of T-shaped stacking is much lower, with a total number of 15.

CONCLUSION

For the buried ionizable residue D271 of Tr. Cel5A, mutating aspartate to alanine or leucine enhances the protein stability. However, for the buried E305, the same mutations decrease the protein stability. QM calculations suggest that the carboxyl stacking interactions with the peptide plane occurring to E305 but not D271 are important for stabilization. Parallel stacking and T-shaped stacking were studied computationally using a complex of model compounds acetic acid (protonated and deprotonated) and NMA. For the protonated carboxyl group, the interaction energy can be as strong as -4 kcal/mol for parallel stacking and -7 kcal/mol for T-shaped stacking. The stronger interaction in the T-shaped conformation comes from a hydrogen bond formed between two compounds. For the deprotonated carboxyl group, the strongest interaction energies for parallel stacking and T-shaped stacking are comparable, about -7 kcal/mol. Due to the solvation effect, the interaction energy can be reduced by as much as about 4 kcal/mol for the deprotonated parallel stacking model and by about 1 kcal/mol for the protonated parallel stacking model. A search of the carboxyl-peptide plane stacking in the PDB databank indicates that parallel stacking is more common than T-shaped stacking, with the most probable distance between the two fragments close to the value predicted by the QM calculations.

ASSOCIATED CONTENT

S Supporting Information

Figure showing the location of D271 and E305 in the CelSA structure (Figure S1). Figure of the partial changes of aspartic acid and glutamic acid side chains (Figure S2). Figure. of dH/d λ for the L271A mutation (Figure S3). Figure of the binding energies in the PCM model (Figure S4). Figure of MM binding energies compared to the QM values (Figure S5). Table of ΔG_3 and ΔG_4 (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: yaols@qibebt.ac.cn. Phone: 86 532 80662792. Fax: 86 532 80662778.

Author Contributions

C. He and J. Chen contributed equally to the work.

Notes

The authors declare the following competing financial interest(s): A Chinese patent has been filed with part of the results in the paper.

ACKNOWLEDGMENTS

This work was supported in part by the 100 Talent Project of Chinese Academy of Sciences, National Natural Science Foundation of China (Grant Nos. 21173247 and 31270785 to L.Y.), Foundation for Outstanding Young Scientist in Shandong Province (Grant No. JQ201104 to L.Y. and ZR2011BQ008 to Y.W.), and "135" Projects Fund of CAS-QIBEBT Director Innovation Foundation.

REFERENCES

- (1) Isom, D. G.; Castaneda, C. A.; Cannon, B. R.; Garcia-Moreno, B. E. Large Shifts in pKa Values of Lysine Residues Buried Inside a Protein. Proc. Natl. Acad. Sci. U.S.A. 2011, 108, 5260–5265.
- (2) Isom, D. G.; Castaneda, C. A.; Velu, P. D.; Garcia-Moreno, B. Charges in the Hydrophobic Interior of Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 16096–16100.
- (3) Trevino, S. R.; Gokulan, K.; Newsom, S.; Thurlkill, R. L.; Shaw, K. L.; Mitkevich, V. A.; Makarov, A. A.; Sacchettini, J. C.; Scholtz, J. M.; Pace, C. N. Asp79 Makes a Large, Unfavorable Contribution to the Stability of RNase Sa. *J. Mol. Biol.* **2005**, *354*, 967–978.
- (4) Bush, J.; Makhatadze, G. I. Statistical Analysis of Protein Structures Suggests that Buried Ionizable Residues in Proteins are Hydrogen Bonded or Form Salt Bridges. *Proteins: Struct., Funct., Bioinf.* **2011**, 79, 2027–2032.
- (5) Kajander, T.; Kahn, P. C.; Passila, S. H.; Cohen, D. C.; Lehtio, L.; Adolfsen, W.; Warwicker, J.; Schell, U.; Goldman, A. Buried Charged Surface in Proteins. *Structure* **2000**, *8*, 1203–1214.
- (6) Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W. Modeling Electrostatic Effects in Proteins. *Bba-Proteins. Proteom.* **2006**, *1764*, 1647–1676.
- (7) Tissot, A. C.; Vuilleumier, S.; Fersht, A. R. Importance of Two Buried Salt Bridges in the Stability and Folding Pathway of Barnase. *Biochemistry* **1996**, *35*, 6786–6794.
- (8) Jayasimha, P.; Shanmuganathan, A.; Suladze, S.; Makhatadze, G. I. Contribution of Buried Aspartic Acid to the Stability of the PDZ2 Protein. *J. Chem. Thermodyn.* **2012**, *52*, 64–68.
- (9) Suominen, P. L.; Mantyla, A. L.; Karhunen, T.; Hakola, S.; Nevalainen, H. High-Frequency One-Step Gene Replacement in *Trichoderma-Reesei* 0.2. Effects of Deletions of Individual Cellulase Genes. *Mol. Gen. Genet.* 1993, 241, 523–530.
- (10) Lee, T. M.; Farrow, M. F.; Arnold, F. H.; Mayo, S. L. A structural Study of *Hypocrea jecorina* CelSA. *Protein Sci.* **2011**, 20, 1935–1940.
- (11) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-balanced, and Scalable Molecular Simulation. *J. Chem. Theory. Comput.* **2008**, *4*, 435–447.
- (12) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, 26, 1701–1718.
- (13) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.

- (14) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Funtions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (15) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y. E. A.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popovic, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P. Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. *Methods Enzymol.* 2011, 487, 545–574.
- (16) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p K_a Predictions. *J. Chem. Theory. Comput.* **2011**, 7, 525–537.
- (17) Case, D. D.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecker, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M-J; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. AMBER 11, University of California: San Francisco, 2011.
- (18) Nose, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50*, 1055–1076.
- (19) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (20) Berendsen, H. J. C. Transport Properties Computed by Linear Response through Weak Coupling to a Bath. In *Computer Simulations in Material Science*; Meyer, M., Pontikis, V., Eds.; Kluwer: Dordrecht, The Netherlands, 1991; pp 139–155.
- (21) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An *N*·log(*N*) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, 98, 10089–10092.
- (22) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (23) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (24) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (25) Simonson, T.; Carlsson, J.; Case, D. A. Proton Binding to Proteins: pK_a Calculations with Explicit and Implicit Solvent Models. *J. Am. Chem. Soc.* **2004**, *126*, 4167–4180.
- (26) Woods, C. J.; Essex, J. W.; King, M. A. The Development of Replica Exchange-Based Free Energy Methods. *J. Phys. Chem. B* **2003**, 107, 13703–13710.
- (27) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, 22, 245–268.
- (28) Song, X. F.; Wang, Y. F.; Shu, Z. Y.; Hong, J. B.; Li, T.; Yao, L. S. Engineering a More Thermostable Blue Light Photo Receptor *Bacillus subtilis* YtvA LOV Domain by a Computer Aided Rational Design Method. *PLoS Comput. Biol.* **2013**, *9*.
- (29) Beveridge, D. L.; Dicapua, F. M. Free Energy Via Molecular Simutation: Application to Chemical and Biomolecular Systems. *Annu. Rev. Biophys.* **1989**, *18*, 431–492.
- (30) Kollman, P. Free Energy Caculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.
- (31) Vangunsteren, W. F.; Mark, A. E. On the Interpretation of Biochemical Data by Molecular Dynamics Computer Simulation. *Eur. J. Biochem.* **1992**, 204, 947–961.
- (32) Straatsma, T. P.; McCammon, J. A. Computational Alchemy. *Annu. Rev. Phys. Chem.* **1992**, 43, 407–435.

- (33) Seeliger, D.; de Groot, B. L. Protein Thermostability Calculations Using Alchemical Free Energy Simulations. *Biophys. J.* **2010**, *98*, 2309–2316.
- (34) Headgordon, M.; Pople, J. A.; Frisch, M. J. Mp2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153*, 503–506.
- (35) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Allaham, M. A.; Shirley, W. A.; Mantzaris, J. A Complete Basis Set Model Chemistry 0.1. The Total Energies of Closed Shell Atoms and Hydrides of the 1st-Row Elements. *J. Chem. Phys.* **1988**, *89*, 2193–2218.
- (36) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (37) Woon, D. E.; Dunning, T. H. Gaussian-Basis Sets for Use in Correlated Molecular Calculations 0.3. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (38) Simon, S.; Duran, M.; Dannenberg, J. J. How does Basis Set Superposition Error Change the Potential Surfaces for Hydrogen Bonded Dimers? *J. Chem. Phys.* **1996**, *105*, 11024–11031.
- (39) Frisch, M. J. T., G, W.; Schlegel, H. B.; Scuseria, G. E. Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, Jr., J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Gaussian 09, Revision B.01; Gaussian, Inc.: Wallingford CT, 2010.
- (40) Lever, M. New Reaction for Colorimetric Determination of Carbohydrates. *Anal. Biochem.* **1972**, 47, 273–279.
- (41) Wu, I.; Arnold, F. H. Engineered Thermostable Fungal Cel6A and Cel7A Cellobiohydrolases Hydrolyze Cellulose Efficiently at Elevated Temperatures. *Biotechnol. Bioeng.* **2013**, *110*, 1874–1883.
- (42) Wang, G. L.; Dunbrack, R. L. PISCES: A Protein Sequence Culling Server. *Bioinformatics* **2003**, *19*, 1589–1591.
- (43) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics Modell.* **1996**, *14*, 33–38.
- (44) Cavallo, L.; Kleinjung, J.; Fraternali, F. POPS: A Fast Algorithm for Solvent Accessible Surface Areas at Atomic and Residue Level. *Nucleic Acids Res.* **2003**, *31*, 3364–3366.
- (45) Harder, M.; Kuhn, B.; Diederich, F. Efficient Stacking on Protein Amide Fragments. *ChemMedChem* **2013**, *8*, 397–404.
- (46) Bendova, L.; Jurecka, P.; Hobza, P.; Vondrasek, J. Model of Peptide Bond–Aromatic Ring Interaction: Correlated ab Initio Quantum Chemical Study. *J. Phys. Chem. B* **2007**, *111*, 9975–9979.
- (47) Imai, Y. N.; Inoue, Y.; Nakanishi, I.; Kitaura, K. Amide $-\pi$ Interactions between Formamide and Benzene. *J. Comput. Chem.* **2009**, 30, 2267–2276.
- (48) James, W. H.; Muller, C. W.; Buchanan, E. G.; Nix, M. G. D.; Guo, L.; Roskop, L.; Gordon, M. S.; Slipchenko, L. V.; Gellman, S. H.; Zwier, T. S. Intramolecular Amide Stacking and Its Competition with Hydrogen Bonding in a Small Foldamer. *J. Am. Chem. Soc.* **2009**, *131*, 14243–14245.
- (49) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (50) Yang, Z. X.; Wang, Z. G.; Tian, X. L.; Xiu, P.; Zhou, R. H. Amino Acid Analogues Bind to Carbon Nanotube via $\pi-\pi$ Interactions: Comparison of Molecular Mechanical and Quantum Mechanical Calculations. *J. Chem. Phys.* **2012**, *136*.
- (51) Banas, P.; Mladek, A.; Otyepka, M.; Zgarbova, M.; Jurecka, P.; Svozil, D.; Lankas, F.; Sponer, J. Can We Accurately Describe the Structure of Adenine Tracts in B-DNA? Reference Quantum Chemical

Computations Reveal Overstabilization of Stacking by Molecular Mechanics. *J. Chem. Theory Comput.* **2012**, *8*, 2448–2460.