Evolutionary Stability of Salt Bridges Hints Its Contribution to Stability of Proteins

Xiaofeng Ban [b], Pratik Lahiri [d], Abhishek S. Dhoble [d], Dan Li [e], Zhengbiao Gu [a,b,c], Caiming Li [b], Li Cheng [b], Yan Hong [b], Zhaofeng Li [a,b,c,*], Bhalerao Kaustubh [d,**]

[a] State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, China
[b] School of Food Science and Technology, Jiangnan University, Wuxi 214122, China
[c] Synergetic Innovation Center of Food Safety and Nutrition, Jiangnan University, Wuxi, Jiangsu 214122, China
[d] Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, IL-61801, USA
[e] The Second Military Medical University, Shanghai, China

ARTICLE INFO

ABSTRACT

The contribution of newly designed salt bridges to protein stabilization remains controversial even today. In order to solve this problem, we investigated salt bridges from two aspects: spatial distribution and evolutionary characteristics of salt bridges. Firstly, we analyzed spatial distribution of salt bridges in proteins, elucidating the basic requirements of forming salt bridges. Then, from an evolutionary point of view, the evolutionary characteristics of salt bridges as well as their neighboring residues were investigated in our study. The results demonstrate that charged residues appear more frequently than other neutral residues at certain positions of sequence even under evolutionary pressure, which are able to form electrostatic interactions that could increase the evolutionary stability of corresponding amino acid regions, enhancing their importance to stability of proteins. As a corollary, we conjectured that the newly designed salt bridges with more contribution to proteins, not only, are qualified spatial distribution of salt bridges, but also, are needed to further increase the evolutionary stability of corresponding amino acid regions. Based on analysis, the 8 mutations were accordingly constructed in the 1,4-α-glucan branching enzyme (EC 2.4.1.18, GBE) from *Geobacillus thermoglucosidans* STB02, of which 7 mutations improved thermostability of GBE. The enhanced thermostability of 7 mutations might be a result of additional salt bridges on residue positions that at least one of amino acids positions is conservative, improving their contribution of stabilization to proteins.

1. Introduction

Salt bridges are defined as electrostatic interactions between two oppositely charged groups: the anionic carboxylate of either glutamate (E) or aspartate (D), and, the cationic ammonium from either arginine (R) or lysine (K) [6,25]. Salt bridge interactions are frequently found in proteins, where they offer structural and functional conformation, and contribute to protein recognition, degradation, catalysis and stability [1,4,8,11,12,32]. From an engineering perspective we ask, 1) is it possible to identify the importance of any particular salt bridge? and 2) can

we engineer salt bridges to predictively alter the biological properties of a protein?

A useful measurement of importance of any particular amino acid within a peptide chain is the degree to which it is conserved. One way of measuring conservation is the evolutionary trace (ET) algorithm [18,22]. The ET is a phylogenetic approach that reliably quantifies the conservation of amino acid residues in proteins, which in turn indicates the evolutionary pressure on the amino acid [2,17,18,21]. It is a powerful tool to identify active sites in enzymes as well residues crucial to protein stability. Generally, amino acid perturbations occurring at positions that are closer to the root of the phylogenetic tree correspond to functional changes, while variations closer to the leaf nodes of the tree correspond to negligible functional changes [27,28,31].

This evolutionary trace approach motivates us to explore the contribution of endogenous salt bridges to proteins from the view of evolution. Previous studies showed that biological properties of proteins depend on cumulatively cooperative interactions of amino acids [9]. In

* Correspondence to: Z. Li, School of Food Science and Technology, Jiangnan University, Wuxi 214122, China.
** Correspondence to: B. Kaustubh, Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, USA.
E-mail addresses: zfli@jiangnan.edu.cn (Z. Li), bhalerao@illinois.edu (B. Kaustubh).

this study we explore not only the conservation of the salt bridge itself, but also the conservation of the local sequence context and its influence on protein stability.

In the present study, we hypothesized that salt bridges influence the evolutionary stability of their neighborhood. The evolutionary stability in our study is different from the term in biological evolution, which was defined as the percentage of residues at a given position that were conserved along the phylogenetic tree. Specifically, altering one residue in a salt bridge pair will not only influence the conservation of the other residue in a salt bridges, but also affect the evolutionary stability of local sequence around the salt bridge. As a corollary, we expected that salt bridges would increase the evolutionary stability of the neighborhood when at least one half of a salt bridge is conserved, which is beneficial for the stability of whole proteins. To validate the hypothesis, we conducted statistical analyses on known salt bridges and attempted to engineer salt bridges in existing structure to positively affect the thermostability of an enzyme.

2. Materials and Methods

2.1. Development of the Salt Bridge Data Set

In order to establish a database of salt bridges, the protein database was selected with resolution no <3.0 Å and an R-factor of no >0.25 [6,12]. For each protein, a set of homologous sequences was retrieved from the NCBI Entrez non-redundant protein sequence database with the E-value <0.05 [7]. In addition, if two sequences shared >99.5% sequence identity, one was arbitrarily selected.

Salt bridges can be defined as an interaction between two oppositely charged groups, where the distance of heavy atom in each pair is the <4.0 Å. Normally, Arg (R) or Lys (K) serve cationic ammonium (RNH_3^+) as positive groups. Simultaneously, Asp (D) or Glu (E) provide anionic carboxylate ($RCOO^-$) as negative groups. His (H) presents an ambiguous protonation state at pH 7.0, which is the reason that we excluded it from our dataset. Therefore, only Arg-Glu, Arg-Asp, Ly-Glu and Lys-Asp salt bridges were considered in this study (The salt bridges used were presented in Appendices). Salt bridges were defined as "local salt bridges" when the sequence separation between the oppositely charged residues was no more than five [6].

2.2. Secondary Structure Elements

The types of secondary structure elements were defined by DSSP, which is a program to standardize secondary structure assignment [33]. The DSSP categorizes secondary structures into eight categories: H = "α-helix"; B = "residue in isolated β-bridge"; E = "extended strand, participates in β ladder"; G = "3-helix (3_{10} helix)"; I = "5 helix (π-helix)"; T = "hydrogen bonded turn"; S = "bend" and R = "coil or other random coil" [10,29,33].

2.3. Evolutionary Stability

The evolutionary stability was defined as the percentage of residues at given positions maintained the same residues along the phylogenetic tree, which was performed by counting the percentage of identical residues at given positions of salt bridges against possible substitutions at these positions along the phylogenetic tree. To calculate Evolutionary Stability sequences in our database were queried using BLAST Tool in NCBI database (https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi). From these results, multiple sequence alignments were generated using ClustalW2 (gap open penalty 10, gap extension penalty 0.05). The Evolutionary Stability of was calculated from these alignments as the percentage of residues identical to the query sequence at the position in the alignment. High evolutionary stability represents this part of salt bridges maintains its original type in the evolutionary. Each query of residues at given positions of salt bridges was obtained from BLAST

Tool in NCBI database (https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi). The sequence alignments were carried out by ClustalW2 (gap open penalty 10, gap extension penalty 0.05) [17,18].

2.4. Directed Evolutionary Stability

At give position of salt bridges, one part of salt bridges might evolve to other residues along the phylogenetic tree, which was defined as "directed evolutionary" in our research. The directed evolutionary stability was calculating the percentage of specific residues evolving into other residues at given positions of salt bridges against all possible substitutions along the phylogenetic tree. All possible substitutions at given positions were retrieved by using BLAST Tool in NCBI database (https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi). Each query of residues at given positions of salt bridges was retrieved with BLAST and aligned with CLUSTALW (gap open penalty 10, gap extension penalty 0.05) [14,20].

2.5. Regionally Evolutionary Stability

Regional evolutionary stability is defined as the average evolutionary stability of a residue and 3 residues upstream and downstream in a protein sequence. The regionally evolutionary stability was measured by counting the ratio of residues around the salt bridges, whose sequence separation were no more than three apart from salt bridges, against possible substitutions at these positions along the phylogenetic tree. According to the residue positions of salt bridges, the regionally evolutionary stability was inspected by the evolutionary stability of regions including plus/minus three sequence order of salt bridges.

2.6. Bacterial Strains, Plasmids and Growth Conditions

The gene encoding GBE from G. thermoglucosidans STB02 was cloned and subsequently expressed in E. coli BL21. The host of E. coli BL21 (DE 3) was used as expression host of GBE. Plasmid pET-20b(+) harboring gbe gene was used for site-mutagenesis and enzyme production. The site-directed mutagenesis of the GBE gene was carried out by using one-step PCR. The oligonucleotides used for mutations of G. thermoglucosidans STB02 GBE are shown in Supplementary Table S1.

2.7. Expression and Purification of Wild Type GBE and its Mutants

Plasmid-carrying strains were grown on LB medium containing 100 μg/mL ampicillin at 37 °C for 12 h. The culture was transferred into LB broth including 100 μg/mL ampicillin and then incubated at 37 °C until the optical density at 600 nm (OD_{600}) reached 1.0–1.5. The medium was then supplemented with 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) to harvest GBE. After 16 h incubation at 20 °C, the crude enzyme was harvested by ultrasonic destruction of the cell walls at 4 °C for 15 min, followed by centrifugation at 10,000 ×g at 4 °C for 20 min. Subsequently, the crude enzyme solution was purified by two chromatographic steps. In the first step of purification, the crude wild type GBE or its mutants was loaded onto a 5-mL HisTrap HP column equilibrated with buffer A (20 mM imidazole, pH 8.0, 500 mM NaCl). The bound GBE or its mutant was eluted with buffer B (500 mM imidazole, pH 7.5, 500 mM NaCl). Then, the protein-containing fractions were collected and loaded onto a 5-mL HiTrap Phenyl HP column (GE Healthcare) with 10 mM Tris-HCl, pH 7.5, containing 0.2 M ammonium sulfate as the starting buffer. After washing the column with 5–10 column volumes of starting buffer, the bound wild type GBE or its mutants was eluted by using 10 mM Tris-HCl, pH 7.5. The homogeneity of each purified GBE was confirmed by SDS-PAGE.

2.8. Measurement of Activity of Wild Type GBE and Its Mutants

The branching activity of wild type GBE and its mutants was determined by the number of α-1,6 linkages introduced by GBE into amylose

(type III, Sigma). Amylose (100 mg/mL) was pre-dissolved in 2 M NaOH and then neutralized with 2 M HCl. The mixture was adjusted to pH 7.0 with 10 mM MOPS, making final concentration of mixture is 5% (w/v). The reaction was initiated by adding an appropriate amount of wild type GBE and its mutants. The amount of newly synthesized α-1,6 linkages was determined by adding 1 mL sodium acetate (1 M, pH 3.5) and 1000 U of isoamylase to 2 mL sample, followed by incubating the mixture at 37 °C for 20 h. One unit is defined as producing 1 mmol of newly synthesized α-1,6 linkages formed per minute at 55 °C [3,13].

2.9. Determination of Thermostability

The resistance to thermal inactivation of GBE and its mutants was measured by incubating purified enzyme in 10 mM Tris-HCl buffer (pH 7.5) at 60 °C and 65 °C. Samples were taken at several time intervals and the residual activities were determined [3].

2.10. Circular Dichroism Analysis

The methods determining circular dichroism (CD) spectra of GBE and its mutants were as described previously, with some modifications [16]. The CD spectra were measured by the range from 190 nm to 250 nm on a MOS-450 CD spectropolarimeter (Biologic, Claix, Charente, France). The purified GBE and its mutant were contained in a 1 mm path length cell, which were determined on 0.1 nm of step resolution and 1 nm/s of scan rate by CD spectra. All CD spectra were performed in triplicates and corrected for the corresponding protein-free sample [5]. Samples were incubated at 65 °C for 10 min and then cooled at 4 °C.

The denaturation curves of wild type and mutant GBEs were measured by monitoring their CD profiles at 220 nm as the temperature was gradually raised from 50 °C to 80 °C.

3. Results

3.1. Distribution Characteristics of Salt Bridges

Our databases of salt bridges successfully recapitulated the previously reported distribution characteristics of sequence separation, secondary structure preferences, C_α-C_α and N-O distances, which are documented in Supplementary Materials.

The sequence separation between the two halves of the salt bridges were examined and illustrated in Fig. 1(A-B). The results show that the distribution of sequence separation between the two halves of the salt bridges is sharply concentrated in a small range, showing a strong propensity for short sequence separation. The salt bridges whose sequence separation <5 was previously termed a "local salt bridge" [6]. The proportion of local salt bridges against total salt bridges is presented in Supplementary Table S2. The results reveal that approximately one third of salt bridges in proteins is local salt bridges. However, even at a small sequence separation, the distribution of salt bridges has particular biases. As indicated in Fig. 1(a-d), the largest distribution of local Arg-Glu salt bridges is at a separation of five residues, while most of local Arg-Asp salt bridges locates at separation of three or four residues. For both types of Lys based salt bridges, most of local salt bridges occurs in four sequence separation. In consist with our research, Donald, J.E reported
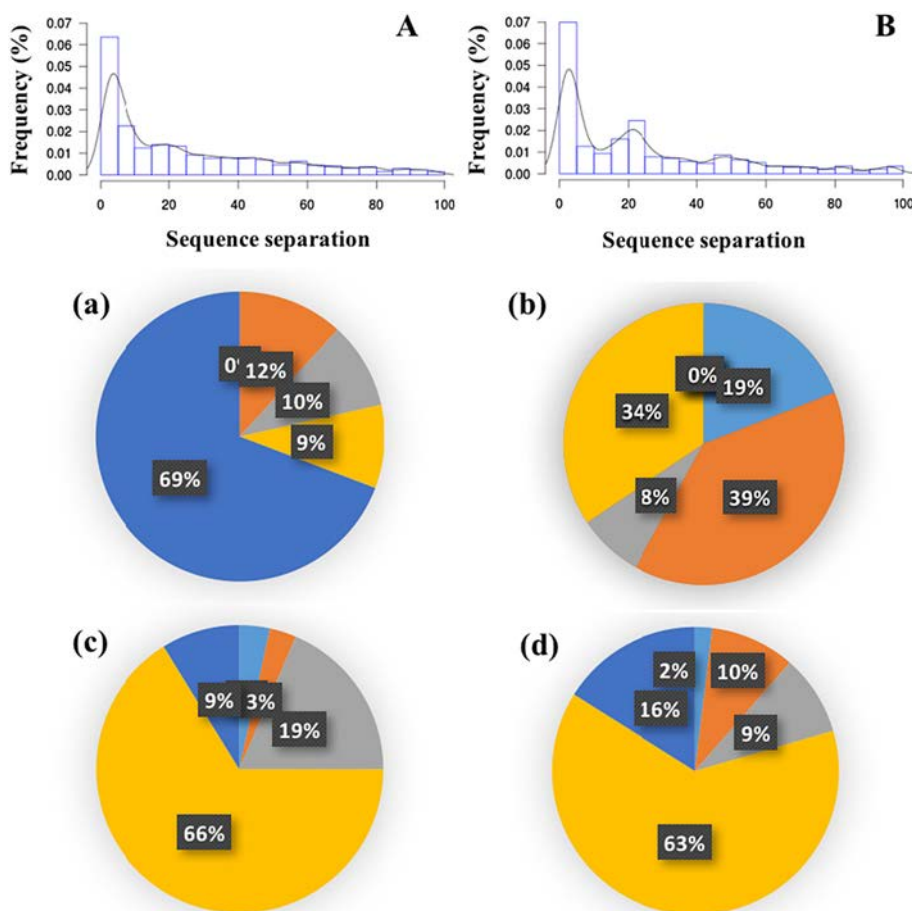
Fig. 1. Distribution of in salt bridges in primary structure. A: Glu based salt bridges, B: Asp based salt bridges, (a): local Arg-Glu salt bridge, (b): local Arg-Asp salt bridge, (c): local Lys-Glu salt bridge, (d): local Lys-Asp salt bridge. The sequence differences, 1: (light blue square), 2: (orange square), 3: (gray square), 4: (yellow square), 5: (dark blue square), between two amino acids in protein sequence.

nearly 34% of Arg, 26% of His, and 31% of Lys salt bridges in the database are local salt bridges [6].

The alpha carbon (C_α) distribution of salt bridges is crucial determinant for predicting the position of salt bridges. Here, the joint distribution between C_α-C_α and N-O distances of the two residues in salt bridges was studied in our case. The results are mapped in Supplementary Fig. S1. A plot of C_α-C_α distance versus N-O distance of Glu and Asp based salt bridges shows that N-O distance concentrates on 2.5 to 4 Å. The C_α-C_α density map of Glu and Asp based salt bridges shows that it dominantly appears at 4–6 Å. However, compared with Glu based salt bridges, the maximum C_α-C_α distance of Asp based salt bridges is 2 Å less than Glu based salt bridges, which might result from shorter length of Asp compared with Glu.

A statistical survey of the secondary structural conformation of salt bridges, which provides insights for the rational design of salt bridges, was investigated in this study and presented in Supplementary Table S3. The results demonstrate that the most dominant secondary structure in Glu and Asp based salt bridges are helix and sheet structures. Additionally, the residues in salt bridges tend to appear at same secondary structures.

The survey of secondary structures in local salt bridges further narrows down the range of rational design of salt bridges. The results are presented in Supplementary Fig. S2. The results show that local salt bridges dominantly appears at sequence separation of three, four and five in helix structure. Analysis of the dataset reveals that the content of secondary structures in Lys based salt bridges is distinct from Arg based salt bridges. However, the two residues in Lys based salt bridges still tend to locate in same types of secondary structures. Our results

also show that Arg and Lys frequently exists in helix structure in local salt bridges, in which Glu tends to locate in helix or coil structures. In contrast, Asp has not shown any preference toward either helix or coil structures.

The survey of distribution characteristics of sequence separation, secondary structure preferences, C_α-C_α and N-O distances of salt bridges lead us the direction of design local salt bridges.

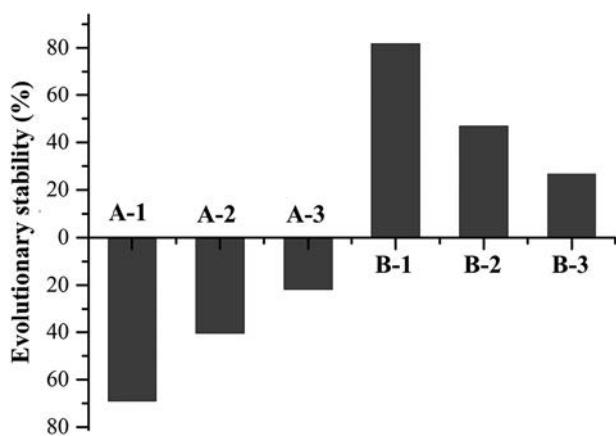### 3.2. Substitutions at Salt Bridge Positions

In this section, we classified salt bridges into two groups based on the identity of the positive residue: Arg based salt bridge, and Lys based salt bridges. The Fig. 2 offers quantitative pictures of the rates of substitutions of salt bridges with other residues along phylogenetic tree. The blue chart showed percentage of possible substitutions occurring at the positions of positively charged residue in salt bridges, while the gray chart presented the rates of possible substitutions appearing at the positions of negatively charged residue in the same type of salt bridge. As shown in Fig. 2, the charged amino acids are most frequently conserved, showing significant evolutionary pressure to conserve these interactions. One exception is the Lys-Glu salt bridge wherein the Glu is substituted to Asp at a small, but significant frequency (~10%).

3.3. Pairwise Evolution at Salt Bridge Positions

We next looked at the pairwise substitutions of salt bridges when one of the positions is conserved. We categorized the alternatives at
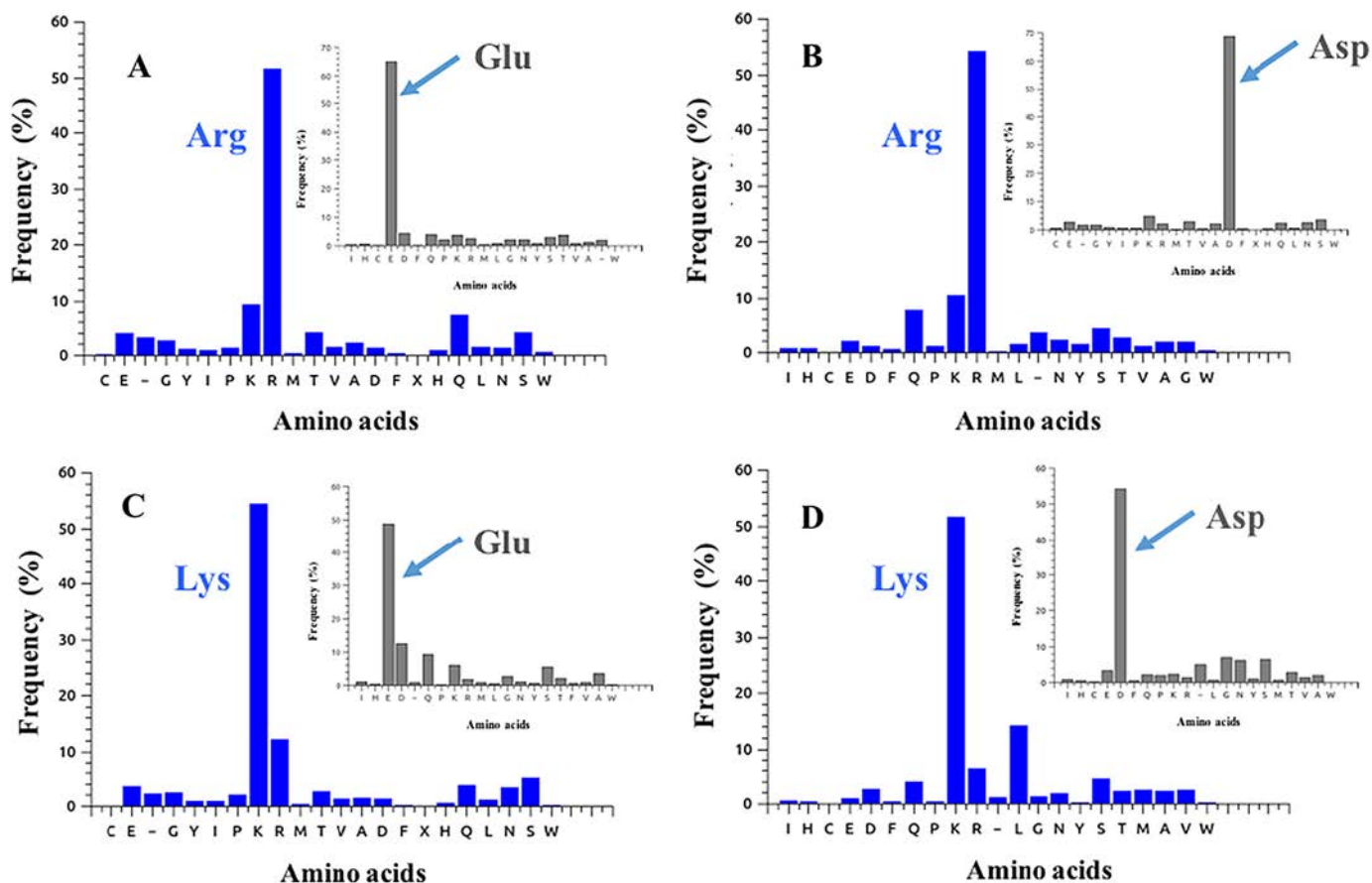
Fig. 2. Frequency of substitutions at given positions of salt bridges. (A) Arg-Glu; (B) Arg-Asp; (C) Lys-Glu; (D) Lys-Asp. The frequency of possible substitutions at, Arg position (A-blue), Glu position (A-gray), in Arg-Glu salt bridge; the frequency of possible substitutions at, Arg position (B-blue), Asp position (B-gray) in Arg-Asp salt bridge; the frequency of possible substitutions at, Lys position (C-blue), Glu position (C-gray), in Lys-Glu salt bridge; the frequency of possible substitutions at, Lys position (D-blue), Asp position (D-gray), in Lys-Asp salt bridge.

Fig. 3. Effects of directed substitutions at salt bridges on evolutionary stability. The vertical axis shows percentage of identical residues at divergent branches. The percentage of identically negative residues (A-1), positive residues (B-1), of salt bridges against total blast results when the other part of salt bridges is conservative; The percentage of identically negative residues (A-2), positive residues (B-2), of salt bridges against total blast results when the other part of salt bridges evolve to other similar residues; The percentage of identically negative residues (A-3), positive residues (B-3), of salt bridges against total blast results when the other part of salt bridges evolve to other neutral residues.

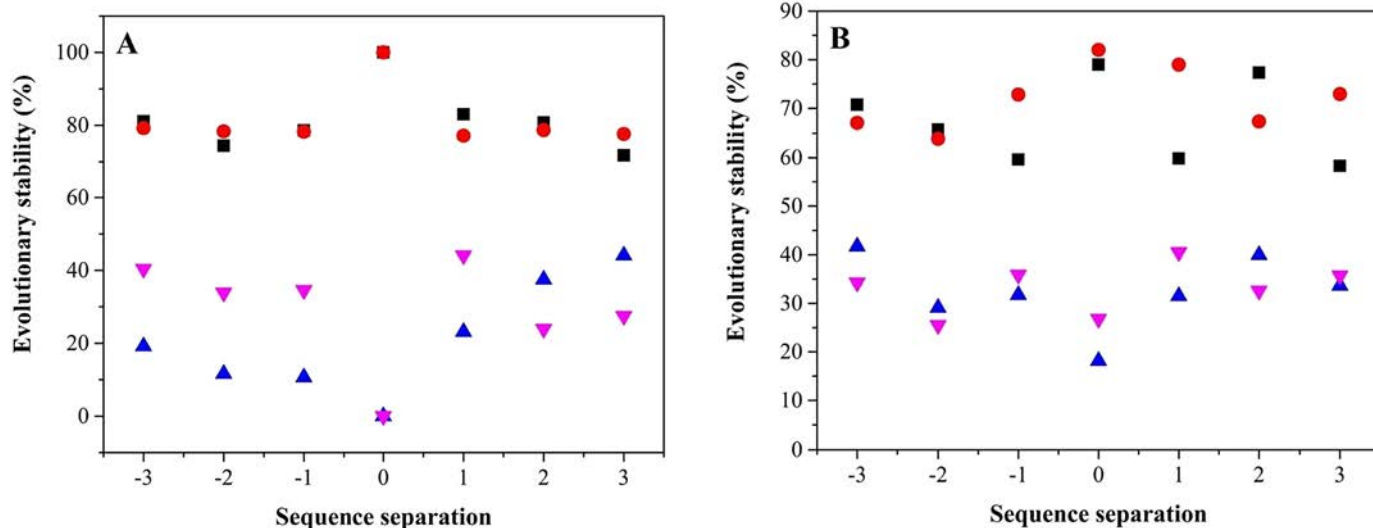### 3.4. Effects of Salt Bridge on Local Evolutionary Stability

Considering that there is a significant co-evolution in salt bridge interactions, we explored the conservation of residues in the local sequence context of salt bridge positions. Since around one-third of all salt bridges are at a small sequence separation (Supplementary Table S2), we restricted the inquiry to 3 residues upstream and downstream of the salt bridge residues. Fig. 4, summarizes the results of this analysis. We see that even when one of the salt bridge residues is conserved, the local sequence around the conserved residue as well as the local sequence around the other salt bridge residue are ~70% conserved each, while this score drops to ~40% when that residue in the salt bridge is not conserved. We can infer that there is correlation between salt bridge conservation and local sequence conservation. Previous results have already shown important differences in the preferred conformation for different types of salt bridges. Hence the results in Fig. 4 point to sequence motifs for different types of salt bridges.

3.5. Relationship of Evolutionary Stability Between Parts of Salt Bridges

We further investigated the effect of salt bridges on their local sequence context via the ET scores which not only measures variations, but, also how close these variations occur to the root of the phylogenetic tree [24,34]. In our study, the results showed that strong correlations of ET scores were observed between positive parts and negative parts of salt bridges (Fig. 5). These results show evidence for the coordinated evolution of the local sequence context around salt bridge residues and hints that alternation of one residue in salt bridge may result in variation occurs in the other corresponding part of salt bridges.

3.6. Rational Design of Salt Bridges

Salt bridges are capable of increasing regionally evolutionary stability of proteins, suggesting electrostatic interactions improve the contribution of amino acids around salt bridges to stability of proteins. Besides, one part of salt bridges could affect the other part of salt bridge in terms of evolutionary stability. According to the analysis of evolutionary characteristics of salt bridges, salt bridges that may have significant contribution to stability are likely to be conservative, or at least one part of salt bridges is conservative.

the second position into three categories- 1) conserved 2) substitution to a charged amino acid of the same polarity 3) substituted to a neutrally charged residue. The results of this analysis are presented in Fig. 3. We see that there is significant evolutionary pressure to maintain a salt bridge (category 1, 2). However, 20% of homologs belonged to category 3, implying that the loss of a charged residue at one of the salt bridge positions does not always lead to loss of the corresponding charged residue. This indicates that charged amino acids are interchangeable to some extent in salt bridges and there is some flexibility in the design of salt bridges. Earlier reports examining the energy of salt bridges through mutation studies have shown that flipping the salt bridge, substituting one charged residue with another of the same polarity can change the energy/thermostability of the bond [15,19]. Thus, salt bridges could part be a mechanism to make fine adjustments to thermostability of proteins.

Fig. 4. Effects of salt bridges on regional evolutionary stability. (A) Regional evolutionary stability of (□), Arg based salt bridges; (○), Lys based salt bridges, when one part of salt bridges was conservative. Regional evolutionary stability of (△), Arg based salt bridges; (▽), Lys based salt bridges, when one part of salt bridges was unconservative. (B) Regionally evolutionary stability of the other part of (□), Arg based salt bridges; (○), Lys based salt bridges, when one part of salt bridges was conservative. Regionally evolutionary stability of (△), Arg based salt bridges; (▽), Lys based salt bridges, when one part of salt bridges was unconservative.
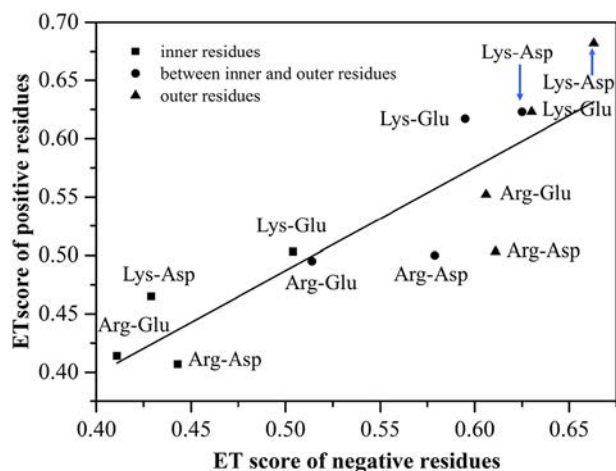
Fig. 5. Relationship of ET scores between residues in salt bridges.

Based on above analysis, salt bridges should be designed following the procedures.

1. The potential salt bridges (salt bridges that will be formed after introducing mutants) should include at least one of endogenously charged residues in order to decrease the blindness of design;
2. The residues involving in potential salt bridges, including endogenous and mutant residues, should locate in the same secondary structures;
3. When sequence separation of two residues is 3 or 4, select the residues that locate in helical structures; when sequence separation of two residues is 2, the residues locating in beta sheet is preferred;
4. The sequence separation of residues in potential salt bridges is <5;
5. The C_α distance between residues is between 4 Å and 14 Å. Simultaneously, the distance between N^+ and O^- atoms of the side chains in potential salt bridges is in the range of from 3.0 Å to 4 Å.
6. The above criterions are apparent properties of salt bridges, which should be strictly qualified when designing salt bridges; the inherent characteristics of salt bridges determine their contribution to the

stability of enzymes. In our selection criteria, the residues in salt bridges should meet the condition that, at least, one part of potential salt bridges is conserved. The conserved residues are defined as their conservation was above 70% by using aligning corresponding sequences with ClustalW2 (https://www.ebi.ac.uk/Tool/msa/clustalw2/).
7. Use PyMOL to visualize and check the possibility of potential salt bridges. Exclude the mutants with residue clash or residues may not form potential salt bridges.
8. The enzyme used in our study, the 1,4-α-glucan branching enzyme (GBE; EC 2.4.1.18) from *G. thermoglucosidans* STB02 (GenBank accession no. KJ660983), is able to react with α-(1,4) and/or α-(1,6) glucosidic linkages and subsequently synthesizes α-1,6-glucosidic bonds, which is a glycoside-transferase belonging to glycosyl hydrolase family 13. Bases on the structure of *G. thermoglucosidans* STB02 GBE, 8 mutants of GBE were constructed by site-directed mutagenesis to test our hypothesis (Supplementary Table S1 and Fig. 6). The characteristics of evolution and structural information of mutant GBE were listed in Table 1. In these mutants, V37E, V37D, Q231R, Q231K, I571D and I266E were built to aim to design local salt bridges. H224E and H224D were unexpectedly selected with above restrictive conditions, however, the H224E and H224D were able to have electrostatic interactions with R170. The C_α distances between two residues are around 4 Å or 6 Å in salt bridges within 5 sequence separation. On the other hand, the C_α distances between salt bridges with >5 sequence separation are out of 4–6 Å.

3.7. Thermostability of Rational Designed Salt Bridges

The mutants were constructed by one-step PCR and verified by DNA sequencing. The wild type and mutant GBEs were expressed in *Escherichia coli* BL21. The results showed that no dramatic differences in expression level between wild type GBE and its mutants were observed. Then, the wild type and mutant GBEs were purified by His Trap and HiTrap affinities chromatography. The homogeneity of each purified GBE was confirmed by SDS-PAGE (Fig. 7).

To determine thermostability, wild type and mutant GBEs were pre-incubated at 60 °C or 65 °C. At the end of the pre-incubation period, residual enzymatic activity was measured using 2.5% soluble starch as
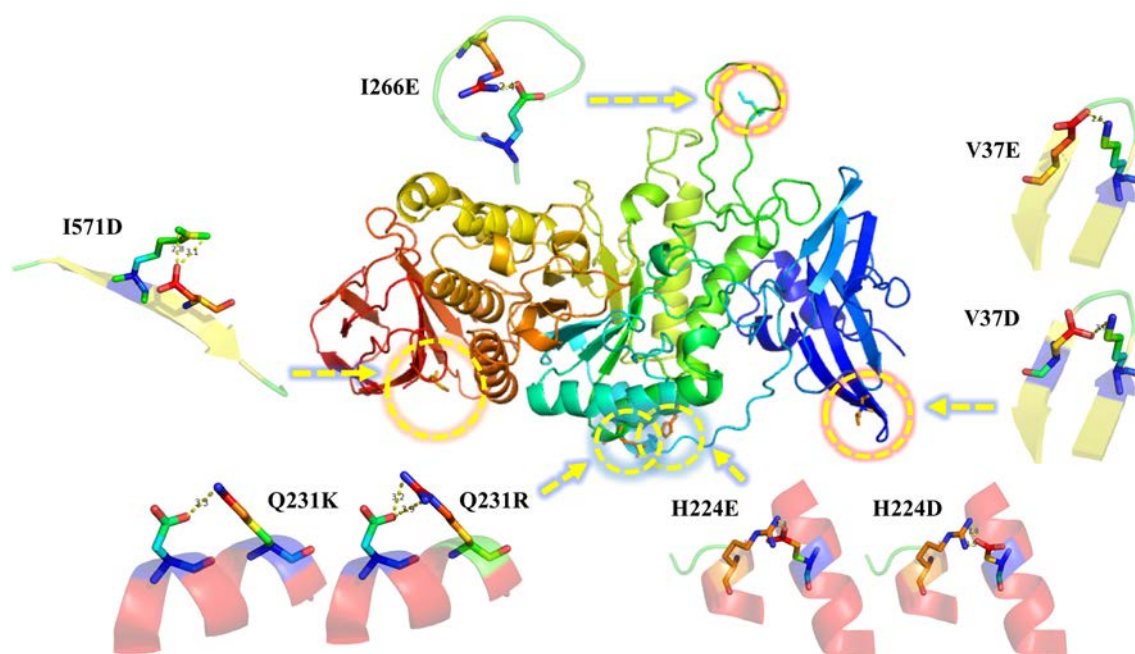

Fig. 6. Schematic presentation of wild type GBE and its mutants.

Table 1
The conservative analysis of possible salt bridges.

Secondary structures of salt bridges	Residue 1[a]	Conservative/Non-conservative[b]	Residue 2[a]	Conservative/Non-conservative[b]	C_α distance (Å)[c]
α-helix	H224	nc	R170	c	6.8
	Q231	nc	D227	c	6.2
β-sheet	V37	c	K32	nc	4.3
	I571	c	R569	c	5.7
Random coil	I266	nc	R269	nc	8.0

[a] The residues of salt bridges.
[b] nc, non-conservative; c, conservatives >70%.
[c] The distance between residue 1 and residue 2.

the substrate. The half-time of wild type and mutant GBEs are presented in Table 2. As indicated in Table 2, the mutant GBEs have similar enzymatic kinetics to wild type GBE, suggesting that the introduced mutations in our study have not significantly affected the catalytic properties of GBE. However, compared with wild type GBE, the half-time (60 °C) of H244D and H244E mutants increases 1.3 and 1.4-folds, respectively. The Q231R and Q231K mutants show substantially increases in thermostability, in which the half-time (60 °C) is improved >1.2 times. The V37E and V37D mutants present modest increases in half-time compared with wild type GBE. The I571D mutant shows approximately 1.5-fold improvement of half-time. In contrast, substitution of Ile with Glu at 266 site results in less effects on thermostability, whose half-time is similar to wild type GBE.

Consistent with thermostability, after incubating at 65 °C for 10 min, the CD spectra analysis show that H244D, H244E, V37E, V37D and I571D have stronger molar ellipticity than wild type GBE, which indicates that these mutants maintain more structural elements compared with GBE after incubation. The results also reveal that these amino acid sites play crucial roles to determine stability of GBE. In contrast to those of above mutants, the CD curve of I266E mutant tends to be flat, suggesting that I266E mutant loss most secondary structures after incubating (Supplementary Fig. S3). To further test the thermostability of mutant GBEs, the T_m values of wild type and mutant GBEs were determined in our research. The results showed that, except I266E, other mutant GBEs with additional salt bridges made the denature curves move to right side of wild type GBE, demonstrating they had higher thermostability compared with wild type GBE.
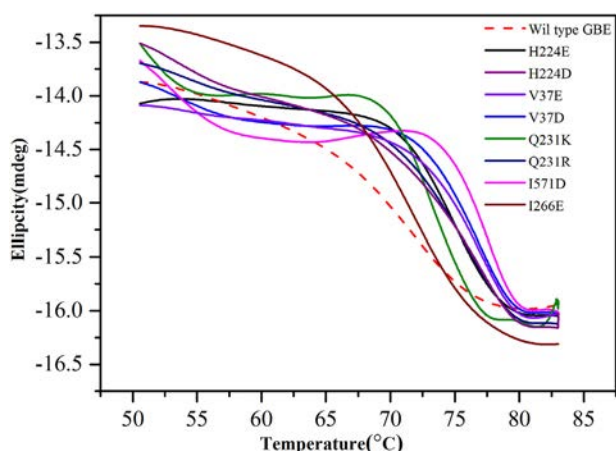
4. Discussion

Salt bridges play important roles in the conformational stability of proteins. However, the contribution of salt bridge has been remained controversial even today [6]. Salt bridges (electrostatic interactions) can be easily formed, however, the contribution of salt bridges to

Fig. 7. The denaturation curves of wild type and mutant GBEs.

stability of proteins varies [9,11]. Several reports have shown less effects of salt bridges on stability, whereas others have demonstrated favorable contributions. Consequently, the rational design of salt bridges is still challenging because of lacking of comprehensive understanding of their contribution to stability of proteins. The aim of this research is to take advantage of the force of nature (Evolutionary trace, ET) to help us to select potential mutants, of which the newly formed salt bridges have significantly positive effects on stability of proteins.

Natural selection is frugal, which eliminates redundant parts of structures in the process of biological evolution [30]. Our study shows that charged amino acids appear more frequently than neutral amino acids at given positions, suggesting that salt bridges are more conservative than non-electrostatic interactions. The inspection of percentage of identical residues against total possible substitutions at given positions of salt bridges was performed at the foundation of database including about 10,000 proteins regardless of species, showing that salt bridges are preserved along the phylogenetic tree, which sufficiently illustrates the importance of salt bridges in biology. The investigation of directed evolution of salt bridges supports that electrostatic interactions tend to be retained along with evolutionary process, demonstrating that salt bridges have stronger evolutionary conservation, which means salt bridges are more vital for stability of proteins [23,26]. The importance of salt bridges rises another question: how salt bridges affect the stability of proteins? The question may not be specifically enlightened, but is able to be generally depicted from the evolution point of view, which provides us a deep insight into the effects of salt bridges on corresponding amino acid regions of proteins. Therefore, the influences of salt bridges on evolutionary stability of corresponding amino acid regions was carried out by counting the ratio of identical residues, whose sequence separation were no more than three apart from salt bridges, against possible substitutions at these positions along the phylogenetic tree. The results reveal that salt bridges are able to increase the evolutionary stability of related amino acid regions. In contrast, breakdown of electrostatic interactions results in decline of regionally evolutionary stability, decreasing their contribution to the whole proteins. The results hint that the new formation of salt bridges not only offers additional electrostatic bonds to corresponding amino acid regions, but also provides the contribution of these amino acid regions to stability of proteins.

What kinds of salt bridges should be built to improve stability of proteins? We followed the two steps of screening. Firstly, the distribution characteristics of sequence separation, secondary structure preferences, C_α-C_α and N-O distances of salt bridges lead the direction of rational design of salt bridges. Secondly, the strong correlation of ET scores between two parts of salt bridges demonstrates that the conservation of residues in salt bridges are tightly correlated (Fig. 5), which can be considered as the electrostatic interactions link ET scores of two parts of salt bridges. In addition, the results of directed evolution of salt bridges show that the conservation of one part of salt bridges is able to affect that of another residue in salt bridges (Fig. 3). Taken together, the newly formed salt bridges should contain at least one of conservative residues, which will increase the conservation of other part of salt bridges. Finally, the new formed salt bridges

Table 2
Activities, K_m values and thermostability of the wild type and mutation GBEs from *G. thermoglucosidans* STB02.

Secondary structures of salt bridges	Mutation	Specific activity (unit/mg)[a]	K_m[a]	Half-life time $t_{1/2}$ (min, 60 °C)[a]	Half-life time $t_{1/2}$ (min, 65 °C)[a]
α-helix	Wild-type	282.3 ± 0.4	1.08 ± 0.02	42.1 ± 1.1	6.9 ± 0.2
	H224D	283 ± 0.2	1.09 ± 0.02	58.2 ± 1.2	8.1 ± 0.3
	H224E	281 ± 0.3	1.08 ± 0.01	59.8 ± 0.8	9.2 ± 0.1
	Q231R	295 ± 0.2	1.01 ± 0.03	53.2 ± 1.7	8.2 ± 0.1
	Q231K	297 ± 0.1	0.98 ± 0.03	53.6 ± 1.9	8.3 ± 0.1
β-sheet	V37E	278 ± 0.2	1.09 ± 0.02	65.2 ± 1.5	10.3 ± 0.0
	V37D	275 ± 0.3	1.12 ± 0.03	69.4 ± 1.7	10.9 ± 0.1
	I571D	280 ± 0.4	1.07 ± 0.01	64.2 ± 1.2	11.3 ± 0.3
Random coil	I266E	280 ± 0.1	1.09 ± 0.04	42.7 ± 1.0	6.9 ± 0.0

[a] Each number presents the mean of three independent measurements.

with conservative residues are beneficial to the evolutionary stability of small amino acid regions, which increase their contribution to stability of whole proteins.

Based on survey of bioinformatics and evolutionary stability of salt bridges, 8 mutations, H244D, H244E, V37E, V37D, I571D and I266E, are constructed according to structural information of *G. thermoglucosidans* STB02 GBE. In the 266 and 571 sites of GBE, only I571D and I266E were constructed. The I571E was excluded because of the residue clash between Glu571 and Arg569, while the I266D was excluded since the distance of heavy atom in each pair is the >4.0 Å. Except for I266E, the rest of 7 mutations, which consist of at least one conservative residues, prolong half-time of GBE compared with that of wild type GBE. These evidences support our hypothesis that salt bridges that are built on at least of one part of conservative residues have more contribution to stability of proteins. Despite a salt bridge between Glu266 and Arg269 strengthens the rigidity of coil region, the mutant I266E has similar activity and thermostability to the wild type GBE, which may result from this mutant occurs in flexible coil regions that has less effects on biological functions of GBE. Interestingly, all the mutants displayed higher activation energy than that of wild type GBE, showing that all the mutants had higher thermostability than wild type (Table S4). Furthermore, design of salt bridges in helix are more stable than design of salt bridges in sheet structures. This may result from the structure of GBE from *G. thermoglucosidans* STB02, which typically contains three domains designated A, B, C. The mutants in residue positions H224 and Q231 are in Domain A, while residue positions I571 and V37 are in Domain C and B. Compared with Domain B and C in GBE, a central (β/α) barrel catalytic domain and essential residues are in Domain A, which in some extent have more effects on performance of GBE.

However, several factors are related with thermostability of proteins, which are not capable of summarizing in one rule. Among these rules that may give characteristics of salt bridges, the evolutionary stability of salt bridges is one aspect of factors that influences on the contribution of salt bridges to stability of whole proteins. The investigation of evolutionary stability of salt bridges also offers us a new way of understanding the contribution of salt bridges. In addition, the evolutionary survey provides us a new approach to rationally design of salt bridges with significant contribution to stability of proteins.

Acknowledgements

Declarations of Competing Interests

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2019.06.022.

References

[1] Albeck S, Unger R, Schreiber G. Evaluation of direct and cooperative contributions towards the strength of buried hydrogen bonds and salt bridges 1. J Mol Biol 2000; 298:503–20.
[2] Amin SR, Erdin S, Ward RM, Lua RC, Lichtarge O. Prediction and experimental validation of enzyme substrate specificity in protein structures. Proc Natl Acad Sci U S A 2013;110:E4195–202.
[3] Ban X, Li C, Gu Z, Bao C, Qiu Y, Yan H, et al. Expression and biochemical characterization of a Thermostable branching enzyme from Geobacillus thermoglucosidans. J Mol Microbiol Biotechnol 2016;26:303.
[4] Bosshard HR, Marti DN, Jelesarov I. Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. J Mol Recognit 2004;17:1–16.
[5] Caiming L, Xiaofeng B, Zhengbiao G, Zhaofeng L. Calcium ion contribution to thermostability of cyclodextrin glycosyltransferase is closely related to calcium-binding site CaIII. J Agric Food Chem 2013;61:8836–41.
[6] Donald JE, Kulp DW, Degrado WF. Salt bridges: geometrically specific, designable interactions. Proteins Struct Funct Bioinformatics 2011;79:898–915.
[7] Donna M, Jim O, Kim DP, Tatiana T. Entrez gene: gene-centered information at NCBI. Nucleic Acids Res 2004;33:D54.
[8] Elcock AH, Mccammon JA. Electrostatic contributions to the stability of halophilic proteins. J Mol Biol 1998;280:731–48.
[9] Gribenko AV, Makhatadze GI. Role of the charge-charge interactions in defining stability and halophilicity of the CspB proteins. J Mol Biol 2007;366:842–56.
[10] Gvritishvili AG, Gribenko AV, Makhatadze GI. Cooperativity of complex salt bridges. Protein Sci 2010;17:1285–90.
[11] Kumar S, Nussinov R. Salt bridge stability in monomeric proteins 1 ☆. J Mol Biol 1999;293:1241–55.
[12] Kumar S, Nussinov R. Close-range electrostatic interactions in proteins. Chembiochem A Eur J Chem Biol 2015;3:604–17.
[13] Kuntal P, Shiva K, Shikha S, Saurabh Kumar G, Mohammad Suhail AH, Eric X, et al. Crystal structure of full-length Mycobacterium tuberculosis H37Rv glycogen branching enzyme: insights of N-terminal beta-sandwich in substrate specificity and enzymatic activity. J Biol Chem 2010;285:20897–903.
[14] Kuo-Bin L. ClustalW-MPI: ClustalW analysis using distributed and parallel computing. Bioinformatics 2003;19:1585–6.
[15] Lee CW, Wang HJ, Hwang JK, Tseng CP. Protein thermal stability enhancement by designing salt bridges: a combined computational and experimental study. Plos One 2014;9:e112751.
[16] Li C, Huang M, Gu Z, Hong Y, Cheng L, Li Z. Nanosilica sol leads to further increase in polyethylene glycol (PEG) 1000-enhanced thermostability of β-cyclodextrin glycosyltransferase from Bacillus circulans. J Agric Food Chem 2014;62:2919–24.
[17] Lua RC, Lichtarge O. PyETV: a PyMOL evolutionary trace viewer to analyze functional site predictions in protein complexes. Bioinformatics 2010;26:2981.
[18] Lua RC, Wilson SJ, Konecki DM, Wilkins AD, Venner E, Morgan DH, et al. UET: a database of evolutionarily-predicted functional determinants of protein sequences that cluster as functional sites in protein structures. Nucleic Acids Res 2016;44:D308–12.
[19] Makhatadze GI, Loladze VV, Ermolenko DN, Chen XF, Thomas ST. Contribution of surface salt bridges to protein stability: guidelines for protein engineering. J Mol Biol 2003;327:1135–48.
[20] Margelevičius M, Venclovas Č. PSI-BLAST-ISS: an intermediate sequence search tool for estimation of the position-specific alignment reliability. BMC Bioinformatics 2005;6:185.
[21] Mihalek I, Res I, Lichtarge O. A family of evolution-entropy hybrid methods for ranking protein residues by importance. J Mol Biol 2004;336:1265–82.
[22] Mihalek I, Res I, Lichtarge O. Evolutionary trace report_maker: a new type of service for comparative analysis of proteins. Bioinformatics 2006;22:1656–7.
[23] Mirny LA, Shakhnovich EI. Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. J Mol Biol 1999; 291:177–96.
[24] Morgan DH, Kristensen DM, Mittelman D, Lichtarge O. ET viewer: an application for predicting and visualizing functional sites in protein structures. Bioinformatics 2006; 22:2049–50.
[25] Musafia B, Buchner V, Arad D. Complex salt bridges in proteins: statistical analysis of structure and function. J Mol Biol 1995;254:761–70.
[26] Nelson ED, Onuchic JN. Proposed mechanism for stability of proteins to evolutionary mutations. Proc Natl Acad Sci U S A 1998;95:10682–6.

[27] Panagiotis K, Amanda K, Stephen Joseph W, Teng-Kuei H, Lua RC, Angela Dawn W, et al. Single nucleotide variations: biological impact and theoretical interpretation. Protein Sci 2015;23:1650–66.

[28] PC N, S H. Predicting deleterious amino acid substitutions. Genome Res 2001;11:863.

[29] Sarakatsannis JN, Duan Y. Statistical characterization of salt bridges in proteins. Proteins Struct Funct Bioinformatics 2010;60:732–9.

[30] Shionyu-Mitsuyama C, Ito Y, Konno A, Miwa Y, Ogawa T, Muramoto K, et al. In vitro evolutionary thermostabilization of congerin II: a limited reproduction of natural protein evolution by artificial selection pressure. J Mol Biol 2005;347: 385–97.

[31] Srinivasan M, Gross AK, Anne P, Meng EC, Wensel TG, Olivier L. Evolutionary trace of G protein-coupled receptors reveals clusters of residues that determine global and class-specific functions. J Biol Chem 2004;279:8126–32.

[32] Takahashi T. Significant role of electrostatic interactions for stabilization of protein assemblies. Adv Biophys 1997;34:41–54.

[33] Wakamura K, Hirokawa K, Orita K. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 2010; 22:2577–637.

[34] Wilkins A, Erdin S, Lua R, Lichtarge O. Evolutionary trace for prediction and redesign of protein functional sites. Methods Mol Biol 2012;819:29–42.