The FEBS Journal

FEBS PRESS science publishing by scientists

Protein thermal denaturation is modulated by central residues in the protein structure network

Valquiria P. Souza, Cecília M. Ikegami, Guilherme M. Arantes and Sandro R. Marana

Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Brazil

Keywords
glycoside hydrolase; protein structural network; residue interaction network; *Spodoptera frugiperda*; β-glucosidase

Correspondence
S. R. Marana, Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, CP 26077, São Paulo, 05513-970 SP, Brazil
Fax: +55 11 3091 2186
Tel: +55 11 3091 8339
E-mail: srmarana@iq.usp.br

(Received 10 August 2015, revised 3 December 2015, accepted 13 January 2016)

doi:10.1111/febs.13659

Network structural analysis, known as residue interaction networks or graphs (RIN or RIG, respectively) or protein structural networks or graphs (PSN or PSG, respectively), comprises a useful tool for detecting important residues for protein function, stability, folding and allostery. In RIN, the tertiary structure is represented by a network in which residues (nodes) are connected by interactions (edges). Such structural networks have consistently presented a few central residues that are important for shortening the pathways linking any two residues in a protein structure. To experimentally demonstrate that central residues effectively participate in protein properties, mutations were directed to seven central residues of the β-glucosidase Sfβgly (β-D-glucoside glucohydrolase; EC 3.2.1.21). These mutations reduced the thermal stability of the enzyme, as evaluated by changes in transition temperature (T_m) and the denaturation rate at 45 °C. Moreover, mutations directed to the vicinity of a central residue also caused significant decreases in the T_m of Sfβgly and clearly increased the unfolding rate constant at 45 °C. However, mutations at noncentral residues or at surrounding residues did not affect the thermal stability of Sfβgly. Therefore, the data reported in the present study suggest that the perturbation of the central residues reduced the stability of the native structure of Sfβgly. These results are in agreement with previous findings showing that networks are robust, whereas attacks on central nodes cause network failure. Finally, the present study demonstrates that central residues underlie the functional properties of proteins.

Introduction

The representation of a protein structure as a graph has been used to express the general properties of the protein based on the interactions among amino acid residues [1–4]. These representations are known as residue interaction networks or graphs (RIN or RIG) [5] or protein structural networks or graphs (PSN or PSG) [6]. The protein tertiary structure is expressed as a network in which nodes correspond to residues and edges represent either covalent or noncovalent interactions [7–10]. Irrespective of the protein folding motif or topology, protein networks have consistently shown a high clustering coefficient (C) and a small path length (L), indicating a high local connectivity. Therefore, a small number of edges or contacts (usually approximately 4) is normally sufficient to connect any pair of residues, even those distantly located in the primary structure. These properties indicate that the protein structure is a small-world network [7,11]. Notably, only a few residues are essential to form short pathways that connect any pair of nodes in the

Abbreviations
C, clustering coefficient; DSF, differential scanning fluorimetry; k_{unfold}, denaturation rate constant; L, average shortest pathway length; PSG, protein structural graph; PSN, protein structural network; RIG, residue interaction graph; RIN, residue interaction network; Sfβgly β-glucosidase from *Spodoptera frugiperda*; T_m, transition temperature.

network. These residues are termed 'central' because they are important for the overall network. If central residues are removed from the protein structural network, a significantly higher number of connections is necessary to establish a pathway between two different residues. These properties are characteristic to proteins, irrespective of their structural group [7,11].

Previous studies have reported that short-range contacts involving residues local in the primary sequence contribute to the small-world property of RIN. Long-range contacts occurring between residues that are close only in the tertiary structure contribute to a scale-free behavior [7,12]. In a scale-free network, central residues are more abundant than in a randomly connected network. The scale-free networks observed in many biological systems [13] have also been demonstrated to be robust and exhibit a high tolerance to random failures, although they do not tolerate direct attacks on central residues [7].

Central residues have been shown to be involved in protein folding [14], to enrich protein dimerization interfaces [11], to form active sites [5] and to be essential for allosteric pathways [15,16]. Consequently, algorithms have been developed to predict the effects of mutations on protein stability using RIN concepts [17]. However, most of these RIN studies were based on databases composed of previously characterized mutant proteins. These mutants were not planned, nor were they found to extensively cover the central residues in any studied protein.

In the present study, we report a systematic investigation of a set of planned mutants of the β-glycosidase GH1 Sfβgly (β-D-glucoside glucohydrolase; EC 3.2.1.21) from the fall armyworm *Spodoptera frugiperda* [18], an $(\alpha/\beta)_8$-fold enzyme that contains replacements of the main central residues and mutations of residues surrounding a main central residue. The Sfβgly mutants were characterized based on their thermal denaturation, which is an appropriate parameter for probing the participation of the central residues in protein stability.

Results and Discussion

The average model structure of Sfβgly obtained from a short molecular dynamics simulation was represented as a network of residues (nodes) linked by interactions (edges) in accordance with previously described procedures [15] (Fig. 1). The network parameters L (average shortest contact pathway) and C (clustering coefficient) for the wild-type Sfβgly were 4.8 and 0.485, respectively. Because L is higher than expected for a random network (2.6) but C is much higher than $C_{random} = 0.02$, the Sfβgly network presumably exhibits 'small-world' properties [7,19], indicating that a small number of contacts (an average of 4.8 for Sfβgly) is sufficient to link any pair of residues [11,20]. In the Sfβgly RIN, we identified only 11 central residues of 509 protein residues (Fig. 2 and Table 1). Interestingly, these central residues are concentrated in the C-terminal half of the Sfβgly structure (10 of 11). Moreover, three of these residues (R97, Y331 and E399) are part of the active site, in accordance with the previous observation that active-site residues tend to be the central nodes of RIN [5].

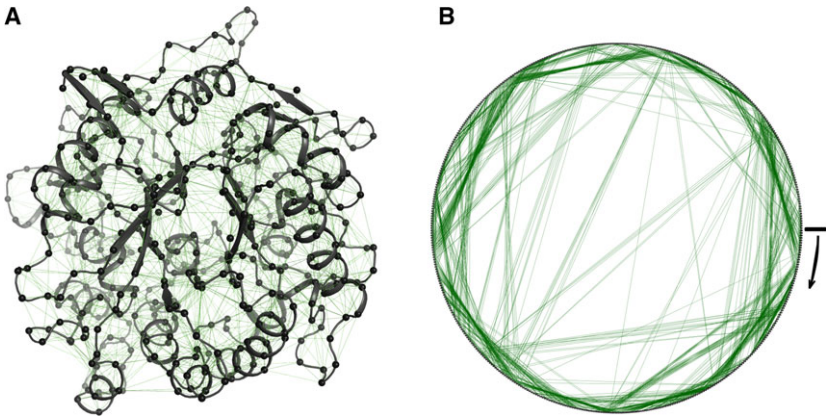

Fig. 1. Residue interaction network of Sfβgly. (A) Homology structure of Sfβgly represented as a network. Interactions presenting a distance ≤ 5 Å indicated by green lines are mapped on the structure. Residues are presented as spheres placed at the C_α coordinates. (B) Circular diagram representation of the Sfβgly residues interaction network. Residues are represented as black dots evenly distributed around the circle in a clockwise order starting from the black bar, which indicates residue 1. Interactions between residues are represented by green lines.

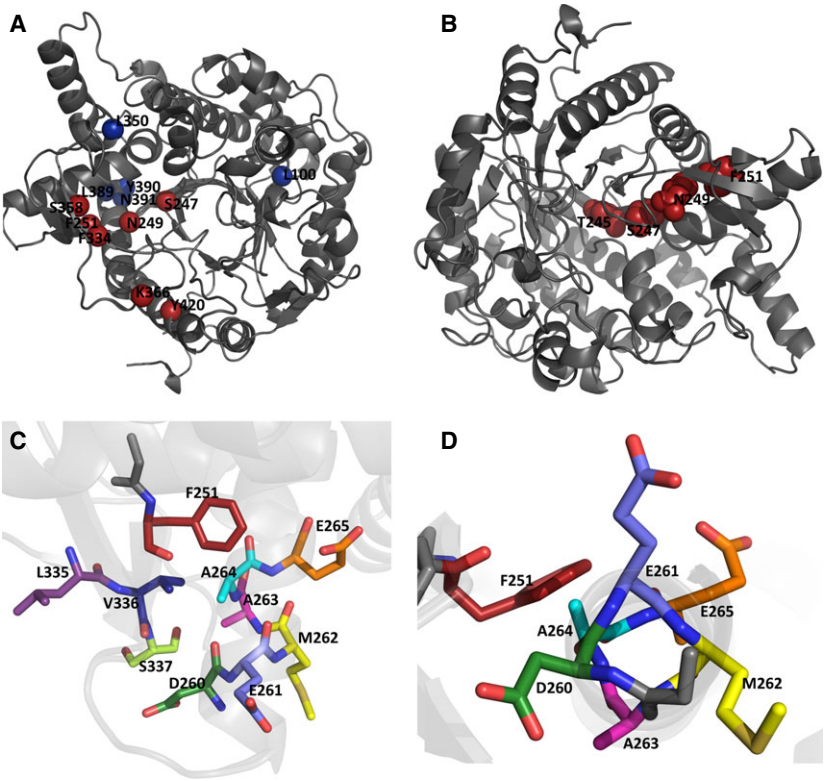

Fig. 2. Spatial distribution of mutation sites on the Sfβgly crystallographic structure (5CG0; this structure will be described elsewhere). (A) Central residues of the Sfβgly structure are presented as red spheres, whereas non-central residues are presented as blue spheres. (B) Set of central residues running from the periphery of Sfβgly to the protein core. (C) Surrounding area of the central residue F251 highlighting the α-helix that contains D260 to E265 and the β-strand that contains L335 to S337. (D) A view parallel to the axis of the α-helix showing that D260, E261 and A264 are located on the α-helix face pointing toward F251.

Table 1. Centrality score for residues of the Sfβgly structure network. *Z*-score = $(\Delta L_k - \Delta L_{k\ average})/\sigma$ [15]. Only residues in which deletion significantly increased *L* (> 2.9 σ) are presented.

Residue	*Z*-score
R97	6.42
F251	6.07
S358	5.11
E399	4.21
T245	4.20
K366	4.08
F334	3.65
S247	3.63
N249	3.58
Y420	3.39
Y331	2.92

To experimentally demonstrate that the central residues affect the functional properties of proteins, seven residues in Sfβgly exhibiting high centrality were replaced to produce Sfβgly mutants (Fig. 2). Replacement with alanine removes most of the side chain contacts, which impairs the participation of the mutated residue as a central residue. The central residues R97, Y331 and E399 were not included in this analysis because they are located at the active site [18,21]. Mutant T245A was produced, although it was not soluble and was retained in inclusion bodies during bacterial expression. For comparison purposes, five non-central residues, L100, L350, L389, Y390 and N391, were also mutated (Fig. 2).

In addition to testing the participation of central residues in the propagation of mutational effects through the protein structure, two sets of residues (D260 to E265 and L335 to S337) close to the central residue F251 (Fig. 2) were also replaced with alanine or phenylalanine. These mutations introduce voids or additional volume in the vicinity of F251, respectively, which perturbs its spatial positioning and contacts. The F251 residue was selected because of its high centrality (Table 1) and because it is part of a group of interacting central residues (T245, S247, N249 and F251). The side chains of this group are positioned on the same face of a β-strand that runs from the periphery of Sfβgly toward its core (Fig. 2). This spatial organization and these connections are suggestive of a set of residues that could propagate conformational modifications through the Sfβgly structural network.

Mutant enzymes were produced as recombinant protein in bacteria, which were then purified. Their folding was then assessed on the basis of CD, tryptophan fluorescence spectra and fluorescence suppression by acrylamide, and these tests produced similar results for the wild-type and mutant Sfβgly. (Figs 3–9 and

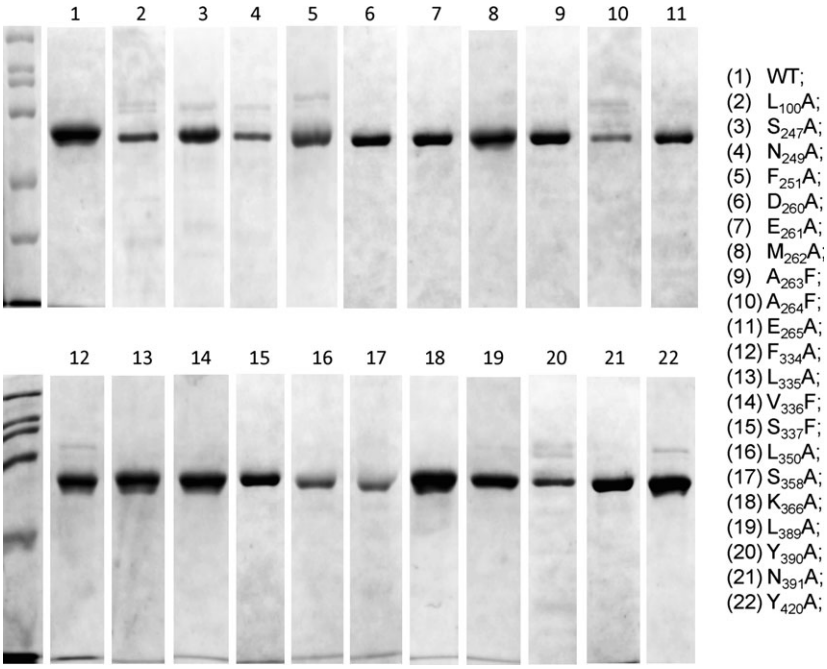

Fig. 3. SDS/PAGE of the purified wild-type and mutant Sfβgly. Gels were 10% polyacrylamide and stained with Coomassie Blue R250.

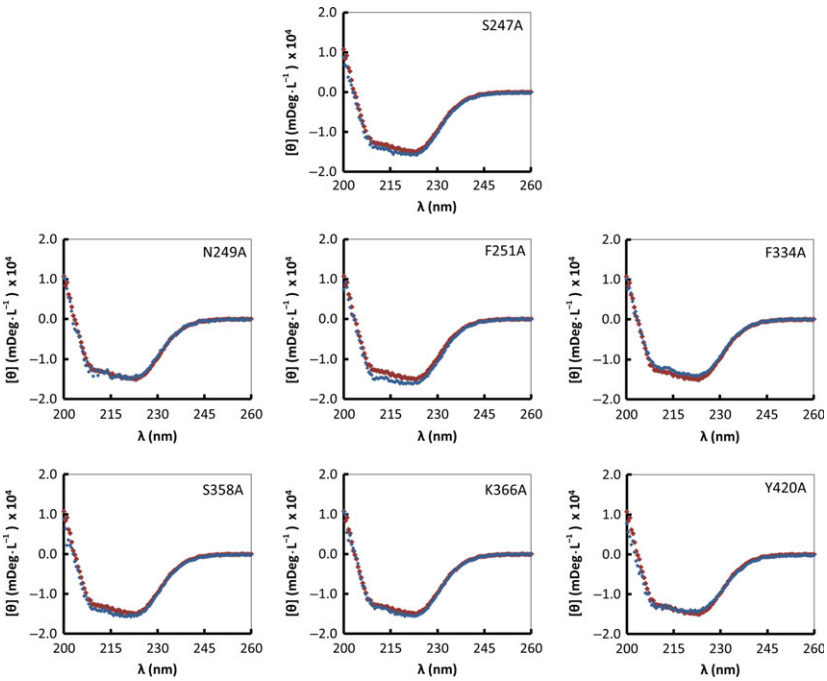

Fig. 4. CD spectra of the wild-type and mutant Sfβgly. Mutations S247A, N249A, F251A, F334A, S358A, K366A and Y420A are replacements of central residues. Blue lines represent mutant Sfβgly; the red line corresponds to wild-type Sfβgly. Spectra were collected at 20 °C in a Jasco 815 spectropolarimeter. Protein samples (0.2 mg·mL^{-1}) were prepared in 10 mM potassium phosphate buffer at pH 6.0.

Table 2). The results obtained suggest that the folding of wild-type and mutant Sfβgly is similar.

The thermal stability of the mutant Sfβgly was then evaluated by determining the T_m and the unfolding rate constant (k_{unfold}) at 45 °C. These parameters reflect the properties of the noncovalent interactions that maintain the protein structure. Thus, they are appropriate for probing the overall behavior of the protein network perturbed by mutations at central residues and their surrounding residues.

The CD spectra and differential scanning fluorimetry (DSF) yielded concordant T_m values (Fig. 10 and Table 3) and revealed a significant decrease in the transition temperature (T_m) for the central residue

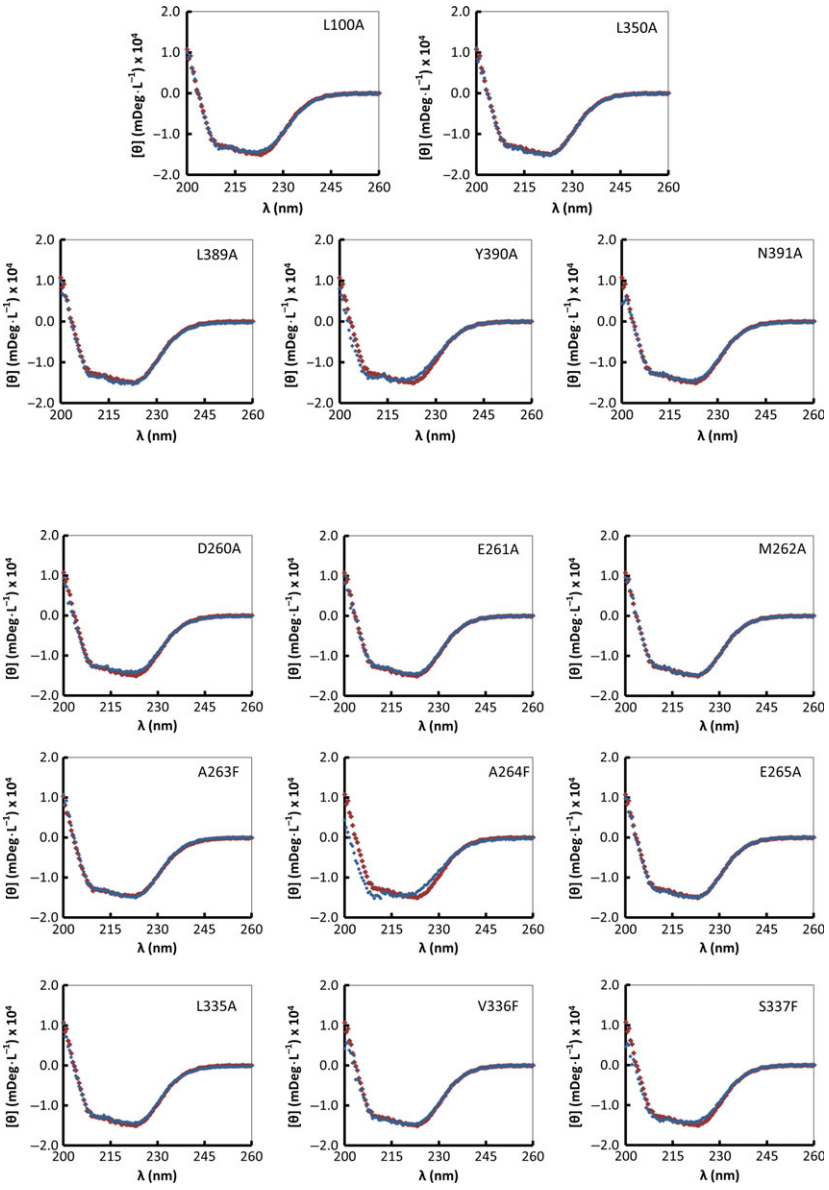

Fig. 5. CD spectra of the wild-type and mutant Sfβgly. Mutations L100A, L350A, L389A, Y390A and N391A are replacements of noncentral residues. Blue lines represent mutant Sfβgly; the red line corresponds to wild-type Sfβgly. Spectra were collected at 20 °C in a Jasco 815 spectropolarimeter. Protein samples (0.2 mg·mL^{-1}) were prepared in 10 mM potassium phosphate buffer at pH 6.0.

Fig. 6. CD spectra of the wild-type and mutant Sfβgly. Mutations D260A, E261A, M262A, A263F, A264F, E265A, L335A, V336F and S337F are replacements of residues in the vicinity of the central residue F251. Blue lines represent mutant Sfβgly; the red line corresponds to wild-type Sfβgly. Spectra were collected at 20 °C in a Jasco 815 spectropolarimeter. Protein samples (0.2 mg·mL^{-1}) were prepared in 10 mM potassium phosphate buffer at pH 6.0.

mutant enzymes N249A, F251A, F334A, S358A, K366A and Y420A. Mutant S247A was an exception, which may be explained by the fact that replacement of serine with alanine is conservative and does not significantly change the network contacts. Notably, mutations of non-central residues (i.e. L100A, L350A, L389A, Y390A and N391A) did not significantly decrease the T_m (Fig. 10 and Table 3). In addition, the k_{unfold} corroborates these observations. Indeed, mutants N249A, F251A, F334A, S358A, K366A and Y420A mutants, in which central residues were replaced, also showed significant increases in the k_{unfold} (Fig. 11 and Table 4). Again, mutant S247A was an exception, as discussed above.

Therefore, in the absence of the contacts provided by central residues, the distribution of conformations visited by Sfβgly has shifted in favor of the less stable conformations, which increased k_{unfold} and decreased T_m. These results are consistent with the observation that thermophilic proteins have more central residues than their mesophilic homologues [22].

An analysis of the mutants designed to probe the participation of the central residue F251 in the propagation of mutational perturbations showed that mutations D260A and A263F caused significant reductions in the T_m (Fig. 10 and Table 3) and effectively increased the k_{unfold} (Fig. 11 and Table 4). The effect of mutation A264F was even more pronounced; it

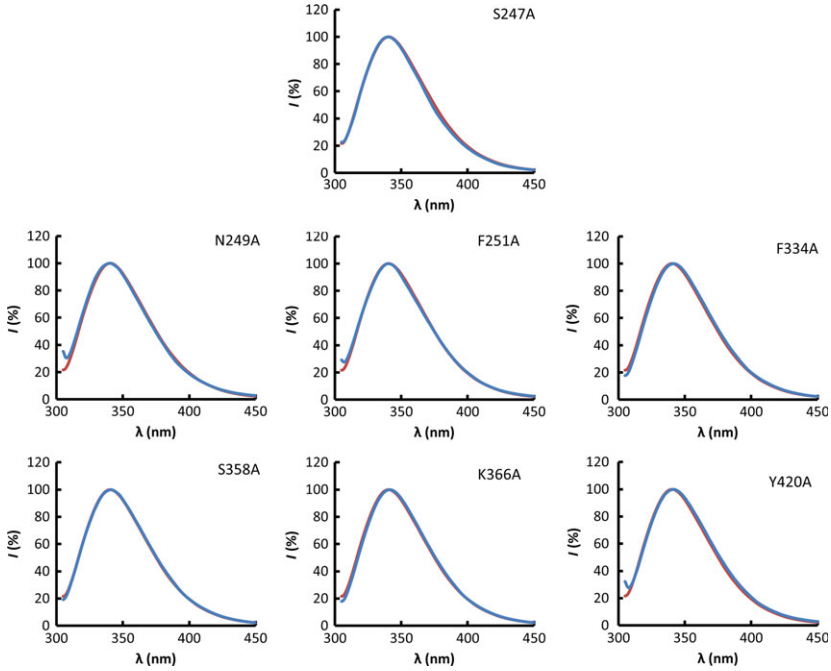

Fig. 7. Tryptophan fluorescence spectra of the wild-type and mutant Sfβgly. Mutations S247A, N249A, F251A, F334A, S358A, K366A and Y420A are replacements of central residues. Blue lines represent mutant Sfβgly; the red line corresponds to wild-type Sfβgly. Spectra were collected at 25 °C in an F-4500 Hitachi spectrofluorimeter. Protein samples were prepared in 10 mM potassium phosphate buffer at pH 6.0.

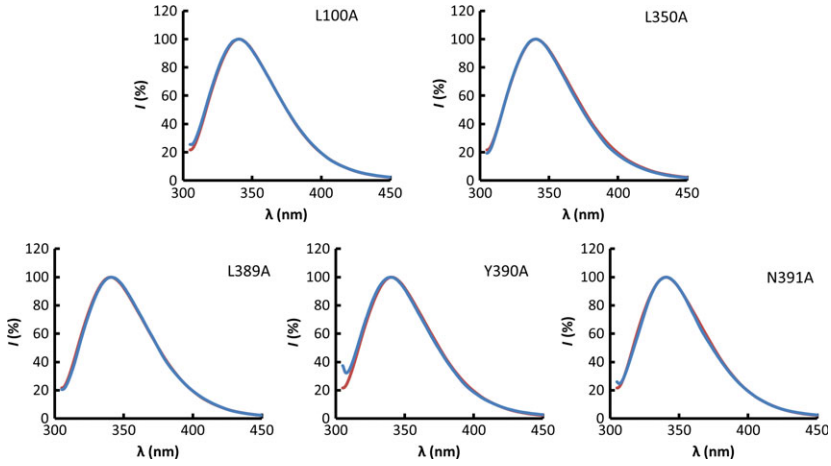

Fig. 8. Tryptophan fluorescence spectra of the wild-type and mutant Sfβgly. Mutations L100A, L350A, L389A, Y390A and N391A are replacements of non-central residues. Blue lines represent mutant Sfβgly; the red line corresponds to wild-type Sfβgly. Spectra were collected at 25 °C in an F-4500 Hitachi spectrofluorimeter. Protein samples were prepared in 10 mM potassium phosphate buffer at pH 6.0.

abolished the two-state thermal denaturation process despite this mutant being folded, as indicated by the CD spectrum (Figs 6 and 12). Conversely, mutations E261A, M262A and E265A, which are also near F251 but located at different positions, did not significantly affect T_m and k_{unfold} (Figs 10 and 11 and Tables 3 and 4). Remarkably, residues 260 to 265 are located in the same α-helix. Therefore, the side chains of D260, A263 and A264 are projected toward F251, whereas those of E261, M262 and E265 are on the opposite face of the helix (Fig. 2). Accordingly, the replacement of A264, in which the side chain points directly toward F251 (contact distance of 3.0 Å) (Fig. 2), caused the most drastic effect.

Because the lack of coupling between domains or discrete structural regions in a protein affects the cooperativity of unfolding [23], the A264F mutant data corroborate the proposal that the perturbation of central residues may disrupt the contacts between discrete clusters of residues over the protein structure [22] and decrease the cooperativity of unfolding. Conversely, the mutation of E265, which is also close to F251 (contact distance of 3.2 Å from the NH group) but has a side chain that points in the opposite direction, did not affect T_m and k_{unfold}. Therefore, despite being restricted to the same α-helix, only changes in side chains pointing toward the central residue F251 significantly affected the thermal stability.

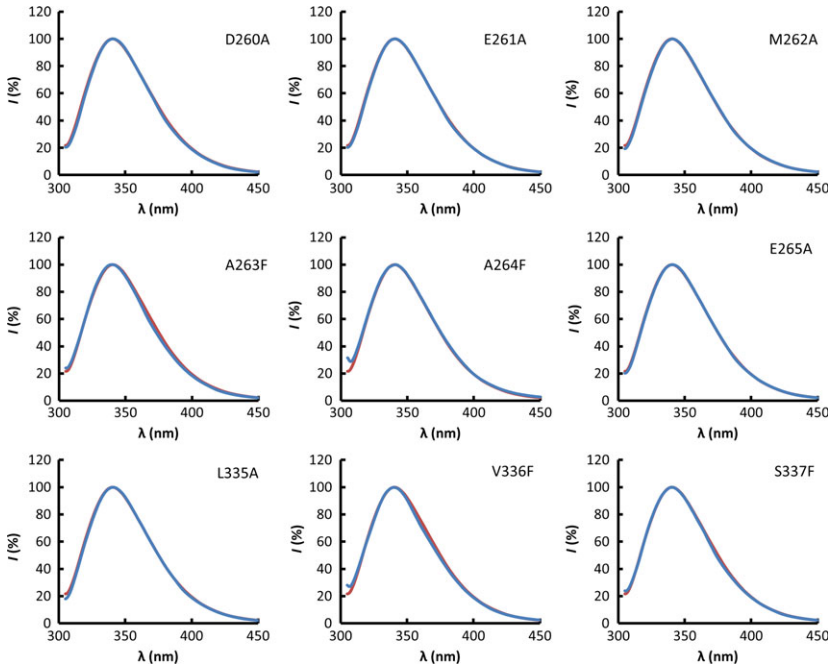

Fig. 9. Tryptophan fluorescence spectra of the wild-type and mutant Sfβgly. Mutations D260A, E261A, M262A, A263F, A264F, E265A, L335A, V336F and S337F are replacements of residues in the vicinity of the central residue F251. Blue lines represent mutant Sfβgly; the red line corresponds to wild-type Sfβgly. Spectra were collected at 25 °C in an F-4500 Hitachi spectrofluorimeter. Protein samples were prepared in 10 mM potassium phosphate buffer at pH 6.0.

Table 2. Stern–Volmer constants (K_{sv}) for acrylamide quenching of the intrinsic fluorescence of the mutant and wild-type Sfβgly. K_{SV} was determined using acrylamide quenching at 25 °C. SDs for the wild-type enzyme were determined based on three different experiments. Central residues are shown in bold. The residues surrounding F251 are in italics; noncentral residues are shown in regular font.

Enzyme	K_{SV}
Wild-type	3.4 ± 0.8
L100A	3.0
L350A	3.4
L389A	3.5
Y390A	2.9
S247A	3.2
N249A	4
F251A	3.6
F334A	3.6
S358A	3.5
K366A	3.3
Y420A	4
D260A	3.0
E261A	3.4
M262A	3.4
A263F	2.8
A264F	3.3
E265A	3.4
L335A	3.4
V336F	3.2
S337F	2.9

Similarly, the second set of mutations near the F251 central residue (i.e. L335A, V366F and S337F) caused significant changes in the T_m and k_{unfold} (Figs 10 and 11 and Tables 3 and 4). Specifically, mutation V336F caused the largest decrease in T_m (approximately 10 °C) and increase in k_{unfold} (by 30-fold). Because residues L335, V336 and S337 are part of the same β-strand, whereas the V336 side chain is projected toward the central residue F251, the side chains of residues L335 and S337 are on the opposite face of the strand (Fig. 2). Accordingly, the relative spatial positioning of the mutated residues to F251 is again decisive with respect to the effects of mutation on thermostability.

The relevance of the central residue F251 in the propagation of mutational perturbation was strengthened by analyzing the thermal stability of a Sfβgly mutant in which residues L389, Y390 and N391 were replaced. In addition to being noncentral, these residues are also located in the immediate area of the non-central residue F280. Specifically, the side chain of Y390 points toward F280, as observed above for pairs A264-F251 and V336-F251. Hence, the L389-N391 residue arrangement is analogous to the L335-S337 set described above, a contiguous group of three residues around a phenylalanine residue. However, all participants here are noncentral. Mutations of L389, Y390 and N391 did not significantly affect T_m and k_{unfold} (Figs 10 and 11 and Tables 3 and 4), confirming that contacts with a central residue are key to the effective propagation of the mutational perturbation through the protein structure.

In conclusion, modifications of the side chains that point toward the central residue F251 were more

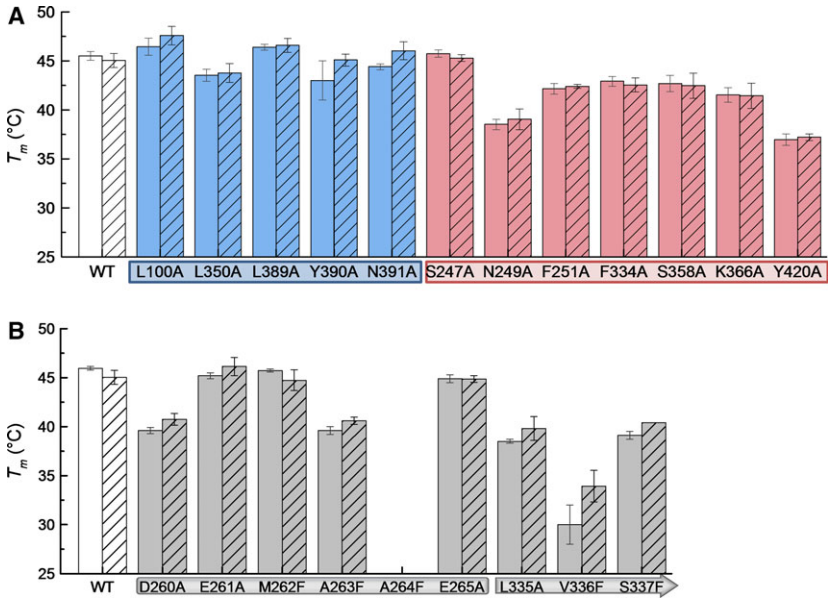

Fig. 10. T_m of the mutant and wild-type Sfβgly. (A) Mutation of central residues (indicated in red) and noncentral residues (indicated in blue). (B) Mutation of residues in the vicinity of F251. Residues D260 to E265, located in an α-helix, are marked with a rectangle, whereas residues L335 to S337, located in a β-sheet, are marked with an arrow. Mutant A264F did not follow a two-state thermal denaturation process, and consequently T_m was not calculated. Plain bars represent data obtained from CD spectra, whereas hatched bars are data obtained from DSF experiments. SEs were calculated based on three independent CD experiments. For DSF, SEs were determined in the fitting processes.

Table 3. T_m of the mutant and wild-type Sfβgly. Central residues are shown in bold. The residues surrounding F251 are in italics; noncentral residues are in regular font. SDs were based on three different experiments. MI indicates that A264F denaturation involves multiple intermediates; hence, there is not a single T_m.

Enzyme	T_m (°C)[a]	T_m (°C)[b]
Wild-type	46.0 ± 0.2	45.0 ± 0.7
L100A	46.4 ± 0.9	47.6 ± 0.9
L350A	43.5 ± 0.6	43.8 ± 0.9
L389A	46.4 ± 0.3	46.6 ± 0.7
Y390A	43 ± 2	45.1 ± 0.6
N391A	44.4 ± 0.3	46.0 ± 0.9
S247A	45.7 ± 0.4	45.3 ± 0.4
N249A	38.5 ± 0.5	39 ± 1[c]
F251A	42.2 ± 0.6	42.4 ± 0.3[c]
F334A	42.9 ± 0.5	42.5 ± 0.7[c]
S358A	42.7 ± 0.8	42 ± 1[c]
K366A	41.5 ± 0.7	41 ± 1[c]
Y420A	37.0 ± 0.6	37.2 ± 0.3[c]
D260A	39.6 ± 0.3	40.8 ± 0.6[c]
E261A	45.2 ± 0.3	46.2 ± 0.9
M262F	45.7 ± 0.2	45 ± 1
A263F	39.6 ± 0.4	40.6 ± 0.4[c]
A264F	MI	MI
E265A	44.9 ± 0.4	44.9 ± 0.3
L335A	38.5 ± 0.2	40 ± 1[c]
V336F	30 ± 2	34 ± 2[c]
S337F	39.1 ± 0.4	40.42 ± 0.01[c]

[a] T_m was determined by monitoring CD readings at different temperatures. SDs were determined in the fitting process.
[b] T_m were determined by DSF.
[c] T_m that is significantly different from the wild-type enzyme ($P \leq 0.01$).

effectively propagated, altering the set of noncovalent interactions that support the entire Sfβgly structure. Moreover, contacts of F251 with the other central residues, such as N249, S247 and T245, may spread perturbations from the mutated external region to the Sfβgly core. This proposition is in agreement with previous studies showing that the allosteric modulation of protein function is mediated by the propagation of 'perturbation waves' through pre-existing pathways interconnected by central residues [15,16,20,24–27]. Similarly, RIN are relevant for structural changes involved in enzyme function [28].

Eyring plots were used to search for details of the mechanism by which the perturbation of central residues change the thermal denaturation of Sfβgly (Figs 13–15). The $\Delta G^{\ddagger}$ values for the denaturation at 45 °C were based on the $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ derived from these plots (Table 5). As expected, the relative sequence of k_{unfold} at 45 °C agrees with the relative ordering of $\Delta G^{\ddagger}$ (Figs 11 and 16). Moreover, the relative sequence of the $\Delta G^{\ddagger}$ also agrees with the ranking of T_m, which was unexpected (Figs 10 and 16). T_m is linked to ΔG^0, which represents the free energy difference between denatured and native Sfβgly, whereas $\Delta G^{\ddagger}$ corresponds to the free energy difference between the transition state of the denaturation process and the native structure. Therefore, the correlation between T_m and $\Delta G^{\ddagger}$ suggests that mutations primarily affect the stability of the native structure of Sfβgly, which is a common determinant of both $\Delta G^{\ddagger}$ and ΔG^0. Thus, the effects of mutations on T_m and k_{unfold} are connected

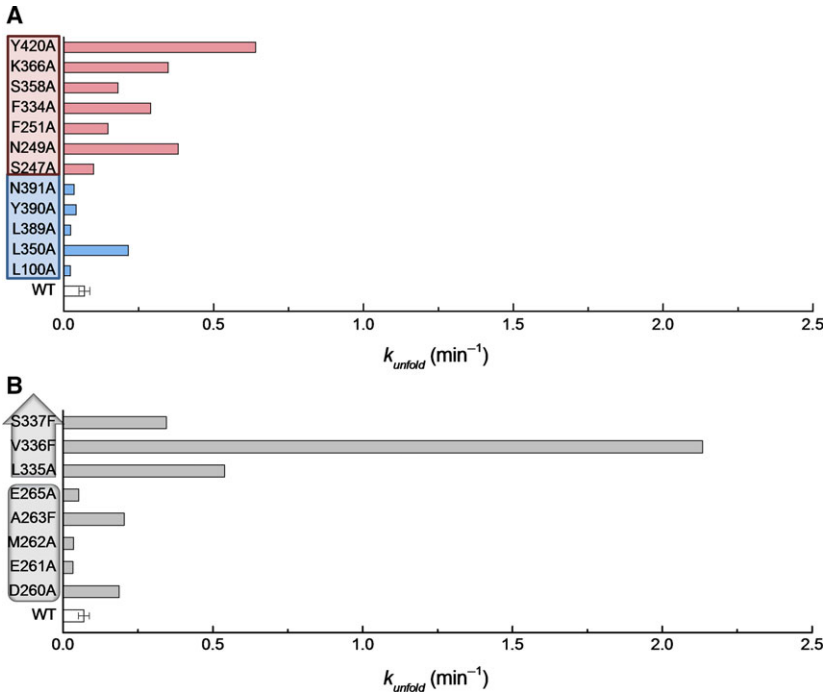

Fig. 11. Rate constants for thermal denaturation at 45 °C (k_{unfold}) of the mutant and wild-type Sfβgly. Data were obtained based on the Sypro Orange binding to the denatured proteins. For the wild-type Sfβgly, the SE was calculated based on seven independent experiments. (A) Mutation of central residues (indicated in red) and noncentral residues (indicated in blue). (B) Mutation of residues in the vicinity of F251. Residues D260 to E265, located in an α-helix, are marked with a rectangle, whereas residues L335 to S337, located in a β-sheet, are marked with an arrow. Mutant A264F did not follow a two-state thermal denaturation process, and consequently k_{unfold} was not calculated.

and have a common cause in the destabilization of the native structure of Sfβgly. In addition, the effects on $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ are notably concerted (Table 5), as observed in other weak interaction systems [29]. For example, both $\Delta\Delta H^{\ddagger}$ and $\Delta\Delta S^{\ddagger}$ were < 0 for Sfβgly mutants that exhibited a clear increase in k_{unfold}, such as V336F and Y420F, and these parameters result in a $\Delta\Delta G^{\ddagger} < 0$ as a result of a compensatory effect. Hence, the $\Delta\Delta H^{\ddagger}$ values agree with the proposition that the deletion or perturbation of central residues reduced the stability of the set of interactions supporting the structure of Sfβgly. Moreover, the entropy increase of the native Sfβgly structure might arise from an initial fragmentation of the structural network in more mobile residues subsets. This mechanism agrees with previous analyses showing that hub deletion causes network failure as a result of fragmentation in partially isolated subsets [30].

In short, mutations of central residues probably altered the set of noncovalent interactions that support the native structure of Sfβgly, likely favoring the less stable conformations of this protein. In addition, mutations of the residues surrounding central residues were more effective in reducing the thermal stability of Sfβgly. Finally, the present study potentially adds to computational analysis and shows that central residues indeed underlie the functional properties of proteins and thus are presumably important for protein thermostability.

Materials and methods

Materials

The vector pCAL-β was constructed as described previously [31]. The PCR reagents were purchased from Thermo Fisher Scientific Inc. (Waltham, MA, USA). The oligos were acquired from IDT (Integrated DNA Technologies Inc., San Jose, CA, USA). The pET-46 Ek-LIC kit was acquired from Merck Millipore (Billerica, MA, USA). The QuikChange Site-Directed Mutagenesis kit and BL21 Gold (DE3) were purchased from Agilent Technologies Inc. (Santa Clara, CA, USA). Sequencing was performed using the BigDye kit from Thermo Fisher Scientific, Inc. Broad range SDS/PAGE standards were purchased from Bio-Rad (Hercules, CA, USA). Sypro Orange dye and other reagents were acquired from Sigma-Aldrich (St Louis, MO, USA).

Sfβgly structural network analysis

The tertiary structure of the β-glycosidase Sfβgly was initially based on the homology modeling in the Swiss Model server [32,33] (depository entry: O61594) taking as template the chain A from the myrosinase of *Brevicoryne brassicae* (1WCG). The homology model was then submitted to a short molecular dynamics simulation (30 ns with the OPLS-AA force field) [34]. Finally, the average model of the tertiary structure of Sfβgly (Dataset S1) determined over the complete simulation was then represented by an

undirected contact network, in which amino acid residues corresponded to nodes, and links are contacts between two residues. A pair of residues was considered in contact if the distance between at least one atom on each residue was smaller than 5 Å, as adopted previously [7]. Residue centrality was computed based on the change of the average shortest pathway length, L, under removal of the residue (node) in accordance with a previous study [15]. Briefly, L was calculated as:

$$L = \frac{1}{N_p}\sum_{j>i} d(i,j)$$

where N_p is the total number of residue pairs (contact distance < 5 Å) and $d(i, j)$ is the shortest network distance between nodes i and j. Contact maps and pathways were calculated with an in-house code, and the shortest network path was found with the Floyd–Warshall algorithm [35]. The average length, L, was recalculated after removal of each residue ($L_{rem.k}$), and the change in the pathway length was determined as: $\Delta L_k = L - L_{rem.k}$. The average $\Delta L_{k,average}$ and its SD (σ) were calculated for the overall residues. Finally, ΔL_k was normalized based on the SD, resulting in the score $z_k = (\Delta L_k - \Delta L_{k,average})/\sigma$. Residues in which deletion increased L by at least 2.9σ were considered central residues. The clustering coefficient (C) of the network was calculated using AMINONET, version 2.0 [36].

Table 4. Rate constants for thermal denaturation at 45 °C of the mutant and wild-type Sfβgly. Central residues are shown in bold. The residues surrounding F251 are in italics; noncentral residues are in regular font. k_{unfold} was determined by monitoring the Sypro Orange binding to unfolded Sfβgly after incubation at 45 °C for different times. The average k_{unfold} value and SD for wild-type enzyme were determined based on seven different experiments. ND, not determined as aresult of native and denatured mutant A264F binding the fluorescent probe similarly.

Enzyme	k_{unfold} (min^{-1})
Wild-type	0.07 ± 0.02
L100A	0.022
L350A	0.22
L389A	0.023
Y390A	0.039
N391A	0.034
S247A	0.10
N249A	0.4[a]
F251A	0.15[a]
F334A	0.29[a]
S358A	0.18[a]
K366A	0.35[a]
Y420A	0.6[a]
D260A	0.18[a]
E261A	0.033
M262A	0.035
A263F	0.20[a]
A264F	ND
E265A	0.05
L335A	0.5[a]
V336F	2.1[a]
S337F	0.34[a]

[a] k_{unfold} values that differed from the wild-type enzyme at the 4σ level.

Expression and purification of the mutant and wild-type Sfβgly

The insert encoding to Sfβgly was amplified by PCR, with pCAL-β as a template and primers 5′-GACGACGAC AAGATCCAGCAGCGCCGCTTCCCCGATGAC-3′ and 5′-GAGGAGAAGCCCGGTTCAATGTCCCTCATCTAT AGTC-3′. pET46Sfβgly was constructed by ligase-independent cloning combining the pET46 Ek/LIC vector and Sfβgly insert in accordance with the manufacturer's instructions. Site-directed mutagenesis was performed with the QuikChange Site-Directed Mutagenesis kit using pET46Sfβgly as template. Site-directed mutations were performed with

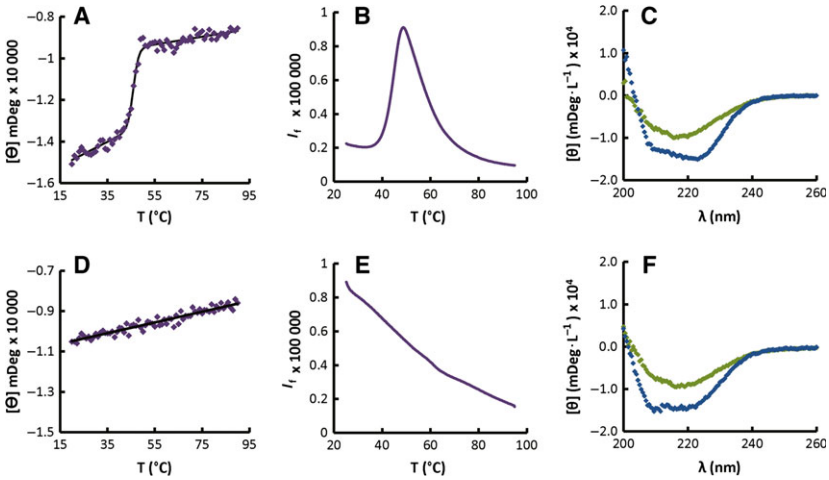

Fig. 12. Thermal unfolding of the wild-type and A264F mutant Sfβgly monitored by CD and DSF. (A) Melting curve of the wild-type Sfβgly based on the CD at 222 nm. (B) Melting curve of the wild-type Sfβgly determined using DSF. (C) CD spectra of the wild-type Sfβgly at 20 °C (blue) and 90 °C (green). (D) Melting curve of the A264F mutant based on the CD at 222 nm. (E) Melting curve of the A264F mutant determined using DSF. (F) CD spectra of the A264F mutant at 20 °C (blue) and 90 °C (green).

the primers reported in Table S1. Mutations were checked by DNA sequencing. BL21 Gold (DE3) competent cells were transformed with wild-type or mutant pET46Sfβgly and plated onto Luria-agar broth containing 100 $\mu g \cdot \mu L^{-1}$ carbenicillin. The cells were grown on 500 mL of Luria-broth containing 100 $\mu g \cdot \mu L^{-1}$ carbenicillin and induced with 1 mM IPTG for 20 h at 20 °C. Subsequently, the cells were harvested by centrifugation and stored at –80 °C. Purification was carried out by suspending a pellet from 100 mL of cultured cells in 2 mL of lysis buffer (50 mM sodium phosphate buffer, pH 8.0, containing 300 mM NaCl and 20 mM imidazole) and sonicating the suspension on ice using five pulses of 10 s at output 3 in a Branson Sonifier 250 (Branson Instruments, Co., Stamford, CT, USA) adapted with a microtip. The lysate was then centrifuged at 11 000 ***g*** for 30 min at 4 °C (5415R centrifuge; Eppendorf; Hamburg, Germany) and the supernatant was collected and then incubated with 400 μL of nickel-nitrilotriacetic acid resin (Qiagen, Valencia, CA, USA) for 14 h with gentle agitation. Subsequently, the resin was washed 10 times with 1.5 mL of lysis buffer, and specifically bound proteins were then eluted with 1.0 mL of 50 mM sodium phosphate buffer at pH 7.0 containing 300 mM NaCl and 300 mM imidazole. Finally, the buffer of the protein sample was changed to 10 mM potassium phosphate at pH 6.0 with a HiTrap Desalting Column (GE HealthCare, Little Chalfont, UK).

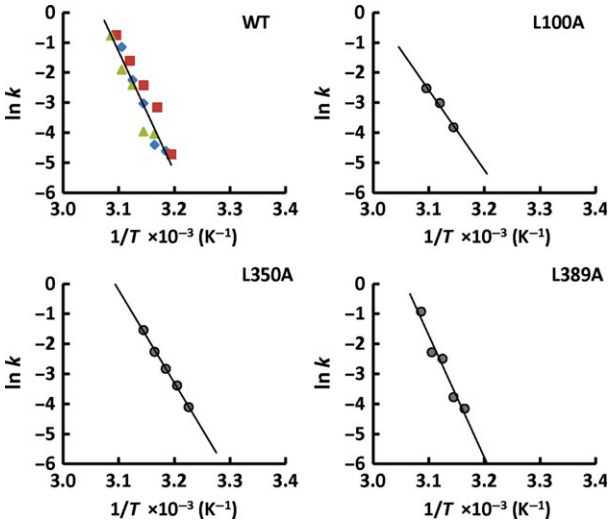

Fig. 13. Eyring plots for the thermal denaturation of the wild-type and mutant Sfβgly Mutations L100A, L350A and L389A are replacements of noncentral residues. k_{unfold} was determined at different temperatures using the same procedure presented in Fig. 11. Three independent experiments were performed for the wild-type Sfβgly (data are presented in different colors).

Determination of the CD spectra

CD was carried out in a Jasco J-815 CD spectropolarimeter (Halifax, NS, Canada) attached to a Peltier system. Data were collected using a rectangular quartz cuvette (NSG Precision Cells Inc., Farmingdale, NY, USA) with a path length of 0.1 cm in continuous scanning mode at a speed

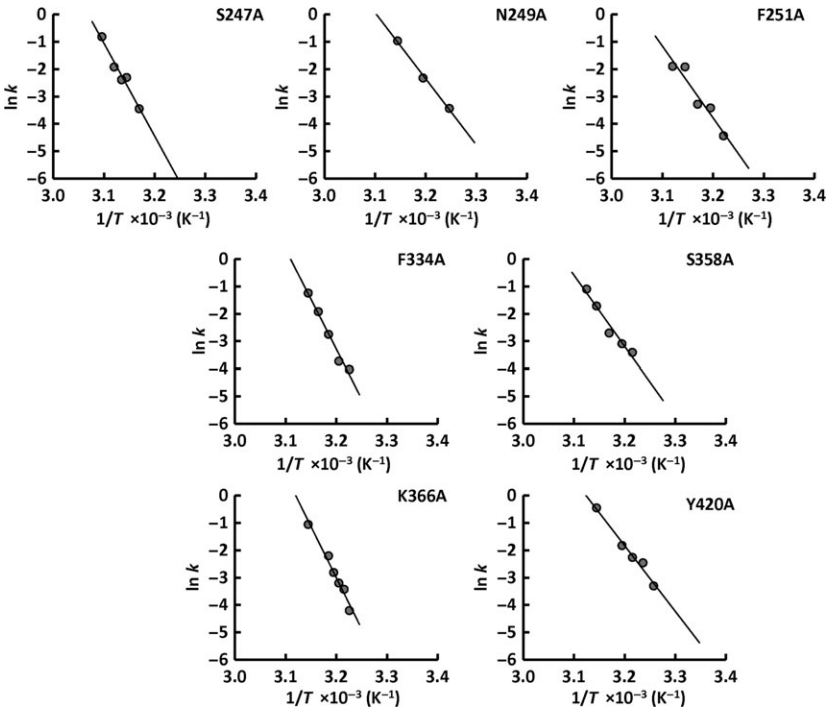

Fig. 14. Eyring plots for the thermal denaturation of the wild-type and mutant Sfβgly Mutations S247A, N249A, F251A, F334A, S358A, K366A and Y420A are replacements of central residues. k_{unfold} was determined at different temperatures using the same procedure presented in Fig. 11.

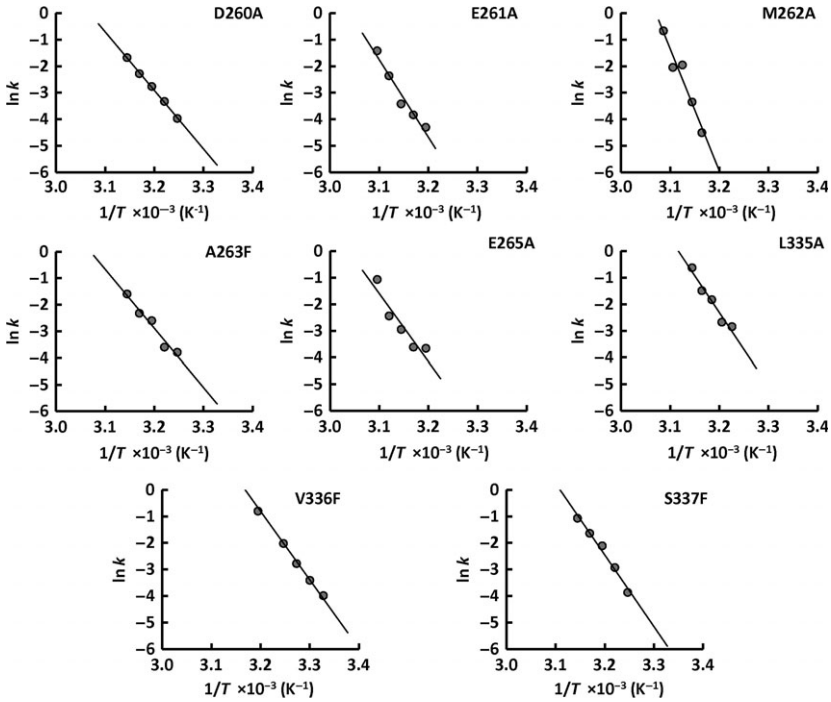

Fig. 15. Eyring plots for the thermal denaturation of the wild-type and mutant Sfβgly Mutations D260A, E261A, M262A, A263F, E265A, L335A, V336F and S337F are replacements of residues in the vicinity of the central residue F251. k_{unfold} was determined at different temperatures using the same procedure presented in Fig. 11.

Table 5. Thermodynamic parameters for thermal denaturation at 45 °C of the mutant and wild-type Sfβgly. Central residues are shown in bold. The residues surrounding F251 are in italics; noncentral residues are shown in regular font. ND, not determined as a result of the A264F mutant showing no transition from the native to the denatured state. *T* was set to 45 °C in the calculations.

Enzyme	$\Delta G^{\ddagger}$ (kJ·mol^{-1})	$\Delta H^{\ddagger}$ (kJ·mol^{-1})	$\Delta\Delta S^{\ddagger}$ (kJ·mol^{-1})	$\Delta\Delta H^{\ddagger}$ (kJ·mol^{-1})	$T\Delta\Delta S^{\ddagger}$ (kJ·mol^{-1})
Wild-type	96.7	355	258	–	–
L100A	100.1	221	121	134	−137
L350A	93.8	256	162	−99	−96
L389A	99.5	338	239	−17	−20
Y390A	95.3	302	207	−53	−52
S247A	96.5	281	184	−74	−74
N249A	92.8	201	108	−154	−150
F251A	93.9	218	124	−137	−134
F334A	92.1	302	210	−53	−48
S358A	94.1	215	121	−140	−137
K366A	91.5	311	219	−44	−39
Y420A	90.8	199	108	−156	−150
D260A	93.9	183	89	−172	−169
E261A	98.9	242	143	−113	−115
M262A	98.9	382	283	27	25
A263F	94.9	184	89	−171	−169
A264F	ND	ND	ND	ND	ND
E265A	98.5	213	114	−142	−144
L335A	93.2	230	137	−125	−121
V336F	87.8	215	127	−140	−131
S337F	93.6	224	130	−131	−128

of 20 nm·min^{-1}, an integration time of 1 s, a bandwidth of 1 nm and standard sensibility. The samples were diluted to 0.2 mg·mL^{-1} (3.5 μM) in 10 mM potassium phosphate buffer at pH 6.0. Eight spectra were collected from 190 to 260 nm and from 20 to 90 °C and averaged. Molar ellipticity was calculated as described previously [37].

Determination of the tryptophan fluorescence spectra

The tryptophan fluorescence spectra of the Sfβgly mutants were collected in a F-4500 Hitachi spectrofluorimeter (Chiyoda, Tokyo, Japan) with a cell holder attached to a circulating water bath. The samples were excited at 295 nm, and three emission spectra were collected from 305 to 450 nm and averaged. The excitation and emission slits were set to 2.5 and 10 nm, respectively. The spectra were measured using a scan speed of 240 nm·min^{-1} at 25 °C.

Tryptophan fluorescence quenching

Quenching was performed by adding aliquots of 10 M acrylamide prepared in 10 mM potassium phosphate buffer at pH 6.0 to the protein solution. The fluorescence intensities were determined at 341 nm with excitation at 295 nm. The dilution of the protein sample was taken into account in the data analysis. Experiments were carried out on a F-4500 Hitachi spectrofluorimeter at 25 °C. The Stern–Volmer constant was calculated using the equation [38]:

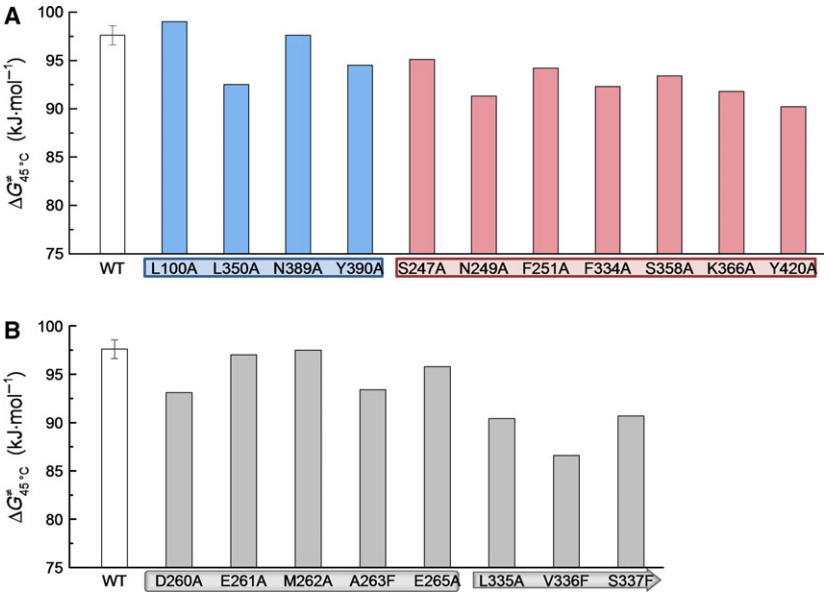

Fig. 16. Free energy of activation ($\Delta G^{\ddagger}$) for the denaturation process at 45 °C of the mutant and wild-type Sfβgly. Data were obtained from the Eyring plots presented in Fig. 11. SEs were determined in three independent experiments performed with the wild-type Sfβgly. (A) Mutation of central residues (indicated in red) and noncentral residues (indicated in blue). (B) Mutation of residues at the vicinity of F251. Residues D260 to E265, located in an α-helix, are marked with a rectangle, whereas residues L335 to S337, located in a β-sheet, are marked with an arrow.

$$\frac{F_0}{F} = 1 + K_{sv}[Q]$$

where F_0 is the fluorescence intensity in absence of acrylamide, F is the fluorescence intensity in the concentration $[Q]$ of acrylamide and K_{sv} is the Stern–Volmer constant.

Determination of the T_m using DSF

DSF [39] was performed on a Real Time PCR 7500 Life Technologies (Carlsbad, CA, USA) instrument. One microgram of wild-type or mutant Sfβgly was mixed with 5000-fold diluted Sypro Orange dye to a total volume of 25 μL in 0.2 mL optical 8 Cap Tube Strips (Life Technologies). Fluorescence data were collected using filter 2 (~ 550 nm) in melting temperature mode at the same time as increasing the temperature from 25 to 95 °C at a rate of 0.1% (~ 2 h of collection). The T_m and SD were determined based on the maximum point of the first derivative of the fluorescence versus T curve in three different experiments.

Determination of the T_m using CD spectra

CD data were collected using a Jasco J-815 spectropolarimeter with a Peltier system. The temperature was increased from 20 to 90 °C at a rate of 0.5 °C·min^{-1}, and each temperature was maintained for 5 s. The samples were diluted to 0.2 mg·mL^{-1} (3.5 μM) in 10 mM potassium phosphate buffer at pH 6.0. The data were collected at three different wavelengths: 222 nm and 208 nm to access the α-helix content and 215 nm to access the β-sheet content, in according with a previous study [37]. To determine the T_m, the data were fitted to the equation:

$$\theta^{\lambda}(T) = \theta^{\lambda}_{U}(T) + \frac{(\theta^{\lambda}_{F}(T) - \theta^{\lambda}_{U}(T))}{1 + \mathrm{Exp}\left(\frac{T - T_m}{a}\right)}$$

where $\theta^{\lambda}_{U}(T)$ is the unfolded protein reading at wavelength λ and temperature T, $\theta^{\lambda}_{F}(T)$ is the folded protein reading also at wavelength λ and temperature T, and a is the curvature parameter. $\theta^{\lambda}_{U}(T)$ and $\theta^{\lambda}_{F}(T)$ correspond to linear equations determined based on experimental data. As a result of sample limitation, T_m for the wild-type and mutant enzymes was determined in a single experiment. The reported SD values are those from the fitting process.

Determination of the rate constant of thermal denaturation

The wild-type and mutant Sfβgly were mixed with 5000-fold diluted Sypro Orange dye in a quartz fluorescence cuvette. The fluorescence data were collected using an F-4500 Hitachi spectrofluorimeter in continuous mode during the incubation at constant temperature. The data were fitted to a first-order kinetics equation to evaluate the rate constant (k_{unfold}):

$$F_t = F_0.e^{-\mathrm{kunfold.t}}$$

where F_t is the fluorescence at time t and F_0 is the initial fluorescence. Seven different experiments were performed for the wild-type enzyme, whereas a single experiment was performed for mutant enzymes. The k_{unfold} values were subsequently plotted and fitted to the Eyring equation:

$$ln(k_{unfold}) = \Delta H^{\ddagger} + \Delta S^{\ddagger}.1/T$$

where k_{unfold} is the rate constant for the protein denaturation at temperature T, and $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ are the transition

enthalpy and entropy, respectively. $\Delta G^{\ddagger}_{T}$ was calculated assuming that $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ are constant over the temperature range considered in the present study.

Acknowledgements

This project was supported by Fapesp (Fundação de Amparo à Pesquisa do Estado de São Paulo; Grants 08/55914-9, 14/21900-2 and 14/19439-5), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), 'Instituto Nacional de Ciências e Tecnologia para o Bioetanol' (Grants: 08/57908-1 and 574002/2008-1) and Capes. We thank Dr Carla Columbano Oliveira for full access to the DSF equipment. SRM is a CNPq research fellow. SRM and GMA are staff members of the Departamento de Bioquímica – IQUSP. VPS and CMI are graduate students supported by Fapesp and Capes, respectively.

Author contributions

SRM planned the experiments. VPS, CMI and GMA performed the experiments. VPS, CMI, GMA and SRM analyzed data. VPS, GMA and SRM were resposible for writing, reviewing and editing. GMA and SRM contributed resources and helped with funding acquisition.

References

1 Csermely P (2008) Creative elements: network-based predictions of active centres in proteins and cellular and social networks. *Trends Biochem Sci* **33**, 569–576.
2 Whitley MJ & Lee AL (2009) Frameworks for understanding long-range intra-protein communication. *Curr Protein Pept Sci* **10**, 116–127.
3 Higman VA, Rosner HI, Ugolini R, Greene LH, Redfield C & Smith LJ (2009) Probing the urea dependence of residual structure in denatured human alpha-lactalbumin. *J Biomol NMR* **45**, 121–131.
4 Boede C, Kovacs IA, Szalay MS, Palotai R, Korcsmaros T & Csermely P (2007) Network analysis of protein dynamics. *FEBS Lett* **581**, 2776–2782.
5 Amitai G, Shemesh A, Sitbon E, Shklar M, Netanely D, Venger I & Pietrokovski S (2004) Network analysis of protein structures identifies functional residues. *J Mol Biol* **344**, 1135–1146.
6 Vishveshwara S, Ghosh A & Hansia P (2009) Intra and inter-molecular communications through protein structure network. *Curr Protein Pept Sci* **10**, 146–160.
7 Greene LH & Higman VA (2003) Uncovering network systems within protein structures. *J Mol Biol* **334**, 781–791.
8 Bagler G & Sinha S (2005) Network properties of protein structures. *Phys a-Stat Mech Applicat* **346**, 27–33.
9 Kundu S (2005) Amino acid network within protein. *Phys a-Stat Mech Applicat* **346**, 104–109.
10 Aftabuddin M & Kundu S (2007) Hydrophobic, hydrophilic, and charged amino acid networks within protein. *Biophys J* **93**, 225–231.
11 del Sol A & O'Meara P (2005) Small-world network approach to identify key residues in protein-protein interaction. *Proteins-Struct Funct Bioinformat* **58**, 672–682.
12 Bartoli L, Fariselli P & Casadio R (2007) The effect of backbone on the small-world properties of protein contact maps. *Phys Biol* **4**, L1–L5.
13 Khanin R & Wit E (2006) How scale-free are biological networks. *J Comput Biol* **13**, 810–818.
14 Vendruscolo M, Dokholyan NV, Paci E & Karplus M (2002) Small-world view of the amino acids that play a key role in protein folding. *Phys Rev E* **65**. doi: 10.1103/PhysRevE.65.061910.
15 del Sol A, Fujihashi H, Amoros D & Nussinov R (2006) Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol Syst Biol* **2**, 0019.
16 del Sol A, Tsai C-J, Ma B & Nussinov R (2009) The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure* **17**, 1042–1050.
17 Giollo M, Martin AJM, Walsh I, Ferrari C & Tosatto SCE (2014) NeEMO: a method using residue interaction networks to improve prediction of protein stability upon mutation. *BMC Genom* **15**, 11.
18 Marana SR, Jacobs-Lorena M, Terra WR & Ferreira C (2001) Amino acid residues involved in substrate binding and catalysis in an insect digestive beta-glycosidase. *Biochim Et Biophys Acta-Protein Struct Mol Enzymol* **1545**, 41–52.
19 Watts DJ & Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442.
20 Atilgan AR, Akan P & Baysal C (2004) Small-world communication of residues and significance for protein dynamics. *Biophys J* **86**, 85–91.
21 Marana SR, Mendonca LMF, Andrade EHP, Terra WR & Ferreira C (2003) The role of residues R97 and Y331 in modulating the pH optimum of an insect beta-glycosidase of family 1. *Eur J Biochem* **270**, 4866–4875.
22 Brinda KV & Vishveshwara S (2005) A network representation of protein structures: implications for protein stability. *Biophys J* **89**, 4159–4170.
23 Shank EA, Cecconi C, Dill JW, Marqusee S & Bustamante C (2010) The folding cooperativity of a protein is controlled by its chain topology. *Nature* **465**, 637-U134.
24 Park K & Kim D (2011) Modeling allosteric signal propagation using protein structure networks. *BMC*

Bioinformatics **12** (Suppl 1), S23. doi: 10.1186/1471-2105-12-S1-S23.
25 Antal MA, Csaba B & Csermely P (2009) Perturbation waves in proteins and protein networks: applications of percolation and game theories in signaling and drug design. *Curr Protein Pept Sci* **10**, 161–172.
26 Goodey NM & Benkovic SJ (2008) Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol* **4**, 474–482.
27 Lockless SW & Ranganathan R (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* **286**, 295–299.
28 Axe JM, Yezdimer EM, O'Rourke KF, Kerstetter NE, You W, Change CA & Boehr DD (2014) Amino acid networks in (b/a)8 barrel enzyme change during catalytic turnover. *J Am Chem Soc* **136**, 6818–6821.
29 Dunitz JD (1995) Win some, lose some – Enthalpy-Entropy compensation in weak intermolecular interactions. *Chem Biol* **2**, 709–712.
30 Albert R, Jeong H & Barabasi AL (2000) Error and attack tolerance of complex networks. *Nature* **406**, 378–382.
31 Marana SR, Terra WR & Ferreira C (2002) The role of amino-acid residues Q39 and E451 in the determination of substrate specificity of the *Spodoptera frugiperda* beta-glycosidase. *Eur J Biochem* **269**, 3705–3714.
32 Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schimidt T, Kiefer F, cassarino TG, Bertoni M, Bordoli L *et al*. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information *Nucleic Acids Res* **42**, W252–W258.
33 Guex N, Peitsch MC & Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30**, S162–S173.
34 Jorgensen WL, Maxwell DS & Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* **118**, 11225–11236.
35 Rosen KH (2002) Discrete Mathematics and Its Applications, 5th edn. McGraw-Hill Publishing Co., New York.
36 Aftabuddin M & Kundu S (2010) AMINONET – a tool to construct and visualize amino acid networks, and to calculate topological parameters. *J Appl Crystallogr* **43**, 367–369.
37 Kelly SM, Jess TJ & Price NC (2005) How to study proteins by circular dichroism. *Biochim Et Biophys Acta-Prot Proteom* **1751**, 119–139.
38 Ghiron CA & Eftink MR (1987) Does the fluorescence quencher acrylamide bind to proteins? *Biochim et Biophys Acta* **916**, 343–349.
39 Niesen FH, Berglund H & Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212–2221.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:
Table S1. Primers for site-directed mutagenesis
Dataset S1. Coordinates of the average structure of Sfβgly (Protein Data Bank file) obtained from the molecular dynamics simulation.