SCIENTIFIC REP**O**RTS

Reversible thermal unfolding of a yfdX protein with chaperone-like activity

Paramita Saha, Camelia Manna, Jaydeb Chakrabarti & Mahua Ghosh

yfdX proteins are ubiquitously present in a large number of virulent bacteria. A member of this family of protein in *E. coli* is known to be up-regulated by the multidrug response regulator. Their abundance in such bacteria suggests some important yet unidentified functional role of this protein. Here, we study the thermal response and stability of yfdX protein STY3178 from *Salmonella* Typhi using circular dichroism, steady state fluorescence, dynamic light scattering and nuclear magnetic resonance experiments. We observe the protein to be stable up to a temperature of 45 °C. It folds back to the native conformation from unfolded state at temperature as high as 80 °C. The kinetic measurements of unfolding and refolding show Arrhenius behavior where the refolding involves less activation energy barrier than that of unfolding. We propose a homology model to understand the stability of the protein. Our molecular dynamic simulation studies on this model structure at high temperature show that the structure of this protein is quite stable. Finally, we report a possible functional role of this protein as a chaperone, capable of preventing DTT induced aggregation of insulin. Our studies will have broader implication in understanding the role of yfdX proteins in bacterial function and virulence.

Complete genome sequencing of several organisms have resulted in new proteins[1] many of which are neither functionally nor structurally characterized. Among such proteins, yfdX proteins are ubiquitously present in virulent gram negative bacteria[2]. Although yfdX protein is identified many years back[3,4], detailed characterization of both function and structure is lacking. The only structural information (PDB 3DZA) on this protein is from *K. pneumoniae* where no function is reported. An earlier study on yfdX protein from *E. coli* reports its up-regulation and co-expression with a regulatory protein evgA, involved in multidrug resistance[3–5]. Our earlier study[2] on yfdX protein SYY3178 from *Salmonella* Typhi reports its interaction with antibiotics namely ciprofloxacin, rifampin and ampicillin. This indicates that yfdX protein might be playing functional role in bacteria.

The stability of protein structure could throw light towards understanding its functionality. The structural stability gets affected by temperature shocks, pH jumps, denaturing agents such as chaotropes, alcohols or acids and results in unfolding[6–10] of the protein. The thermal response for different proteins is diverse. There are many which can tolerate high temperatures, many of which unfold at elevated temperatures but refolds back to the native structure upon cooling[11–20] and other form irreversible aggregates upon heating. For example, polymerases like Taq[21] or pfu[22] from thermophillic bacteria show very high temperature stability. Another class of protein, known as the heat shock proteins (Hsp), like IbpB[23] from *E. coli*, plant cytosolic small Hsp of class I and II[24], GroES[25], to name only a few, are produced at elevated temperatures. They not only possess chaperone activity but also show thermal reversibility in unfolding. Likewise, reversibility in thermal unfolding is shown in by other chaperones, like Skp[26] from periplasm or SlyD[27] and SecB[28] from cytosol of *E. coli*.

Here, we study the thermal responses of the yfdX family protein STY3178. The primary objective of this work is to shed light on its function by characterizing thermal unfolding-folding and stability. Our experimental studies unravel reversible thermal unfolding of this protein. We show the thermal stability of the protein at elevated temperatures from circular dichroism (CD), steady state fluorescence and nuclear magnetic resonance (NMR), whereas the state of oligomerization remains invariant as observed from DLS. We also measure the kinetics of thermal unfolding and refolding from CD. A model of STY3178 based on the homology and molecular dynamics (MD) simulation is proposed to understand the observed thermal stability. Finally, we identify the capability of STY3178 to act as chaperone *in vitro*.

Department of Chemical, Biological And Macromolecular Sciences, S. N. Bose National Centre for Basic Sciences, Block JD, Sector III, Salt lake, Kolkata 700106, India. Correspondence and requests for materials should be addressed to M.G. (email: mahuaghosh@bose.res.in or mahua.ghosh@gmail.com)
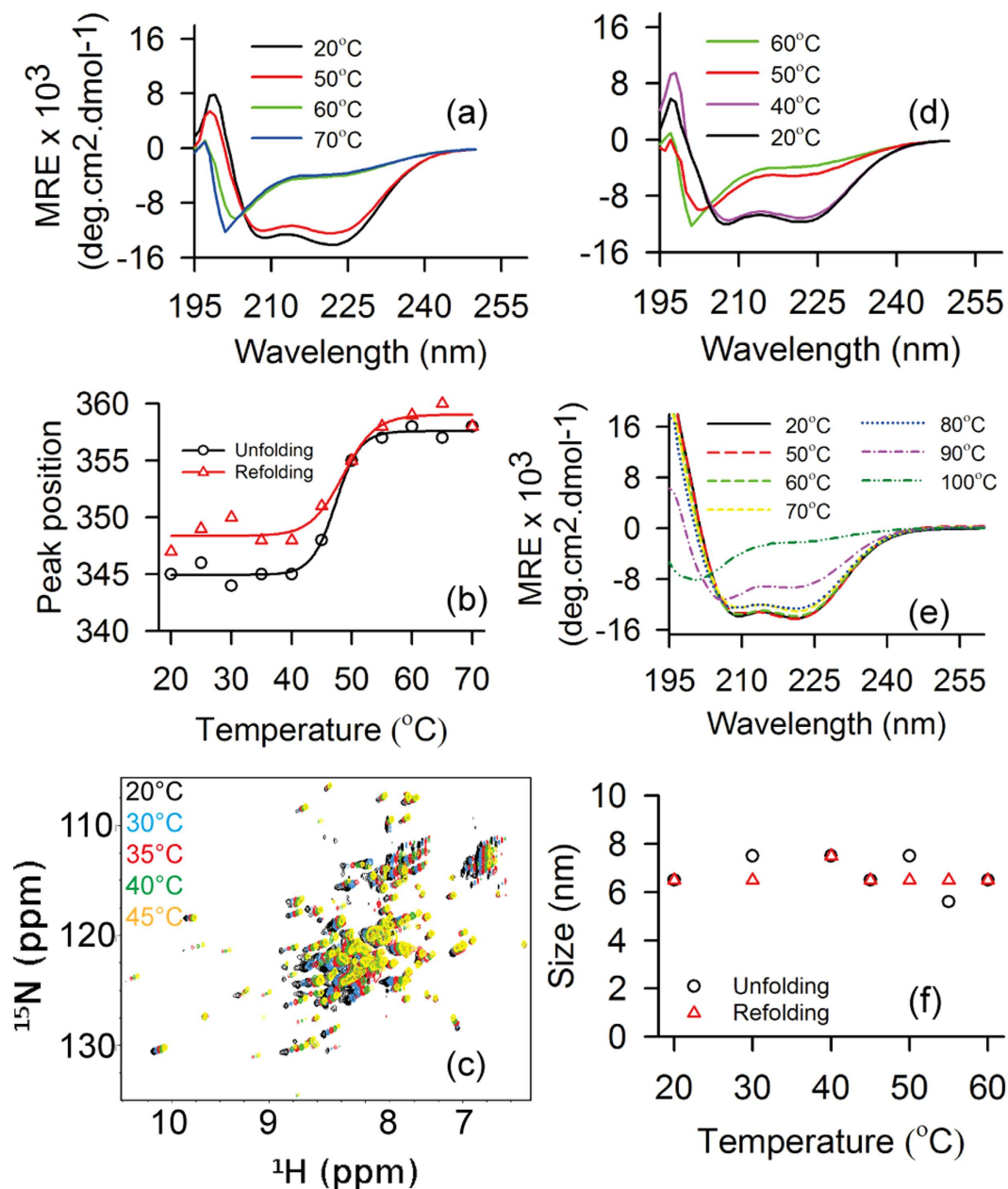
Figure 1. Structural changes of STY3178 at higher temperature. (**a**) Shows the CD spectra of protein for temperatures 20 °C (black), 50 °C (red), 60 °C (green) and 70 °C (blue) during heating. (**b**) Shows the change in fluorescence emission maxima for 280 nm excitation with temperature during unfolding (black) and refolding (red). (**c**) Shows the ^{15}N-^{1}H HSQC spectra (6–10.5 ppm) of the protein for temperatures 25 °C (black), 30 °C (blue), 35 °C (red), 40 °C (green) and 45 °C (yellow). (**d**) Shows the CD spectra upon cooling (refolding) at temperatures 60 °C (green), 50 °C (red), 40 °C (magenta) and 20 °C (black). (**e**) Far UV-CD spectra of the refolded protein recorded at room temperature after heating it for 30 minutes at 50 °C (red), 60 °C (green), 70 °C (yellow), 80 °C (blue), 90 °C (magenta) and 100 °C (dark green). (**f**) Hydrodynamic diameters of STY3178 measured from DLS plotted for temperatures 20 °C, 30 °C, 40 °C, 50 °C and 60 °C during heating (black) and cooling (red).

Results

Circular dichroism (CD) spectrum indicates that STY3178 is an α-helical protein containing two minima around 209 and 222 nm (Fig. 1a, black). The steady state fluorescence emission maximum of this protein is around 342 nm for excitation wavelengths of 275, 280 and 295 nm. This is in agreement to our earlier observation[2]. We probe the thermal unfolding, refolding upon cooling and conformational stability of STY3178 using CD, fluorescence, DLS and NMR.

Temperature induced unfolding and conformational stability. The α-helical secondary structure remains stable and nearly unchanged up to 50 °C when heated gradually in steps of 10 °C starting from 20 °C (Fig. 1a). The structural signature significantly changes beyond 50 °C. The two minima at 209 and 222 nm disappear completely at 60 °C (Fig. 1a) indicating total loss of secondary structure. No further change in the CD spectrum is observed at 70 °C (Fig. 1a) when compared to 60 °C spectrum (Fig. 1a). Thus we find the protein to unfold completely at 60 °C.

Next we monitor the emission maxima position for excitation at 280 nm in steady state fluorescence. The emission peak position of the protein remains unaltered till 45 °C (Fig. 1b), beyond which the maxima starts shifting towards higher wavelength. A jump of nearly 10 nm in the emission peak position is observed between 45 °C to 55 °C. This shift towards higher wavelength is an indication of protein unfolding. Beyond 55 °C, the emission peak position does not change any further. We observe a decrease in fluorescence emission intensity with increase in temperature. These observations are qualitatively similar to those from CD measurement.

HSQC spectrum of ^{15}N-labelled protein shows well dispersed peaks between 6–10.5 ppm at 25 °C (Fig. 1c). We observe large number of peaks clustered around 7.5–8.5 ppm region as found for α-helical proteins in the HSQC spectrum (Fig. 1c). The spectrum also contains well dispersed peaks other than the clusters observed. This indicates presence of some β-sheet elements in STY3178. The dispersion of the HSQC spectrum changes marginally when we increase temperature from 25 °C to 45 °C. This suggests that the tertiary structural fold is indeed stable at 45 °C, in agreement with CD and fluorescence data. There are only a few peaks in the HSQC spectra which show sensitivity upon increasing temperature (Fig. 1c). Interestingly most of the peaks showing sensitivity to higher temperature belong to the well dispersed regions and not the core helical region.

Reversibility in unfolding. We check the reversibility in unfolding by monitoring the refolding using CD by cooling down the protein from an elevated temperature. The thermally unfolded protein at 70 °C is gradually cooled to 20 °C by decreasing temperature in step of 10 °C. We observe the protein remains unfolded up to 50 °C (Fig. 1d). We observe refolding of the protein where the α-helical secondary structure is regained upon further cooling down to 40 °C (Fig. 1d). The CD spectrum after cooling down to 20 °C (Fig. 1d) is very similar to the native protein spectrum (Fig. 1a) prior to thermal unfolding. This observation demonstrates reversibility in thermal unfolding of STY3178. We observe a blue shift ~8 nm compared to the unfolded state in the fluorescence emission peak position upon cooling (Fig. 1b). This again indicates protein refolding where a native-like structure is formed which results in blue shift compared to the unfolded protein. We also observe an increase in fluorescence intensity upon refolding in contrast to the decrease during unfolding. However, the final emission peak intensity is less and the position is red shifted by 2–3 nm in the refolded protein compared to the native, suggesting slight rearrangement in side chains conformations. Overall both steady state fluorescence and CD results are in agreement indicating reversibility in thermal unfolding of this protein.

The conformational stability under equilibrium condition is probed by heating the protein for a longer period of time. We heat the protein at different high temperature water baths (50 °C to 100 °C) for 30 minutes and cool down it to room temperature after that. The CD spectra of the cooled proteins are then compared with the native one. We find that protein when heated within the range 50 °C to 80 °C, refolds completely as native-like structure, shown in Fig. 1e. There is a slight decrease in ellipticity of the refolded protein when it is treated at 90 °C, indicating partial loss of structure (Fig. 1e). We further observe that no refolding could be achieved when the protein is heated at 100 °C (Fig. 1e). Thus, the reversibility of folding is maintained in STY3178 fully when treated at 80 °C.

We measure the hydrodynamic size of STY3178 in the temperature range 20 °C to 60 °C (Fig. 1f) using DLS. The folded protein in solution at 20 °C has a hydrodynamic diameter around 6.5 nm as reported earlier[2]. The variation of this size with temperature is only around 1 nm which is within the fluctuation limit. This indicates that the oligomerization state of the protein does not change upon heating and remains stable. When we cool the same protein from 60 °C to 20 °C, we again observe similar hydrodynamic sizes (Fig. 1f).

Heating or cooling rate dependence. We perform ellipticity measurements for different heating rates. The fraction of folded protein (f_N) is estimated from the ellipticity at 222 nm [θ_{222}] for each temperature (as described in method section). For none of the heating rates below 50 °C, protein unfolds and f_N remains constant. Above 50 °C, f_N decreases for various rates (Fig. 2a). We observe a difference in half denaturation temperature (T_m) that is at $f_N = 0.5$ for the different heating rates. T_m is low for slower heating rates and increases for faster heating rate as tabulated in Table 1. In other words, longer the protein remains at elevated temperature above 50 °C, faster is the unfolding. We also monitor the refolding of the protein using different cooling rates. The plot of f_N versus temperature for different cooling rates (Fig. 2b) show refolding of the protein starts when the temperature is below 50 °C. The half renaturation temperature (T_m') of the protein is thus lower than the half denaturation temperature (T_m) (Table 1), indicating hysteresis in the unfolding and refolding. These observations indicate that the insufficient time during faster scan in temperature hinders the protein to come to equilibrium at a given temperature causing the hysteresis. This is also indicative of kinetically controlled process[29–31]. The hysteresis decreases with decreasing rates of heating and cooling.

Kinetics of unfolding and refolding. The kinetics of unfolding of STY3178 is measured from f_N at a given temperature with time in the temperature range 53 °C to 65 °C (see methods) where the f_N values tend to zero after sufficiently long time (Fig. 3a). f_N decreases below 50% within approximately 50 minutes when the protein is heated in this temperature range. The plot of f_N with time follows single exponential decay. We obtain the rate of unfolding (k_u) from the exponential fit of f_N versus time (few representative cases shown in Fig. 3b). We observe that k_u values follow the Arrhenius behavior as can be seen from lnk_u versus 1/T plot, where T is the absolute temperature (Fig. 3c). The activation energy of unfolding (E_a^u) is calculated from the slope of the plot. Here
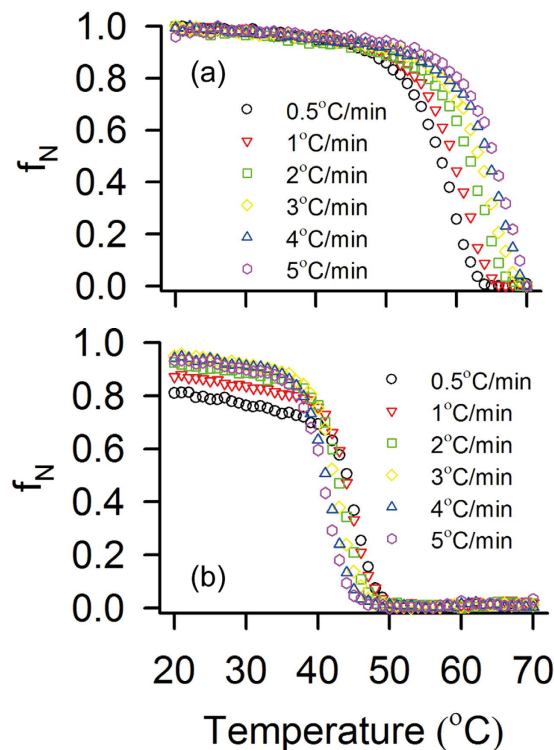
Figure 2. Unfolding and refolding of STY3178. (**a**) Thermal unfolding of STY3178 showing the variation of f_N with temperature for different heating rates 0.5 °C/min (black), 1 °C/min (red), 2 °C/min (green), 3 °C/min (yellow), 4 °C/min (blue) and 5 °C/min (magenta). (**b**) Change in f_N upon cooling the protein from 70 °C at different cooling rates 0.5 °C/min (black), 1 °C/min (red), 2 °C/min (green), 3 °C/min (yellow), 4 °C/min (blue) and 5 °C/min (magenta).

Rate (°C/min)	T_m (°C)	$T_m{}'$ (°C)
0.5	56.57	46.57
1	58.65	45.73
2	61.15	44.59
3	62.4	43.96
4	63.65	43.24
5	64.59	42.4

Table 1. Half denaturation (T_m) and renaturation temperature ($T_m{}'$) of protein STY3178 for different rates of heating and cooling.

$E_a^u \sim 246.9\,\text{kJ/mol}$ obtained for STY3178 is comparable to the activation of energy of unfolding reported in literature[32–34]. In the temperature range below 53 °C the protein does not reach complete unfolded state despite long measurement time. Hence this region has not been considered for estimating E_a^u.

We monitor the refolding kinetics of the protein from the change in ellipticity at 222 nm $[\theta_{222}]$ starting from the thermally unfolded state at 70 °C. The unfolded protein is then directly cooled down to different temperatures in the range of 40 °C to 30 °C. There is a time lag (t_L) of about 100 seconds after which f_N increases rapidly to reach the folded state ($f_N \approx 0.95$) within approximately 300 seconds. The data for the entire time of measurement (0 to 600 seconds), averaged over repeated set of experiments, are shown in Fig. 3d for different temperatures. The finite t_L is probably due to stabilization after temperature quenching and depends slightly on temperature. We calculate $\Delta f_N = f_N - f_R$, where f_R is the folded fraction at $t = t_L$. The f_R values are typically small but non zero for different temperatures. Similarly, we define the rise time $\Delta t = t - t_L$. The plot of Δf_N versus Δt show single exponential rise, few of which are shown in Fig. 3e. We calculate the rate of refolding (k_f), from the fitted curves. Figure 3f shows $\ln k_f$ versus $1/T$ plot confirming the Arrhenius behavior upon refolding. The activation energy (E_a^f) of refolding is $\sim -58.66\,\text{kJ/mol}$.

The quality of linear fit in refolding kinetics (Fig. 3f) is somewhat poorer than that in unfolding (Fig. 3c). This is reflected in the R^2 values of the fits (0.97 for unfolding and 0.85 for refolding). This leads us to check the dependences of $\ln(k_u/T)$ and $\ln(k_f/T)$ on $1/T$ as suggested in ref. 20. The plots, shown in the Supplementary Fig. S1, confirm linear dependences for both the cases. This suggests that the activation heat capacities[20] for both unfolding and refolding are negligible.
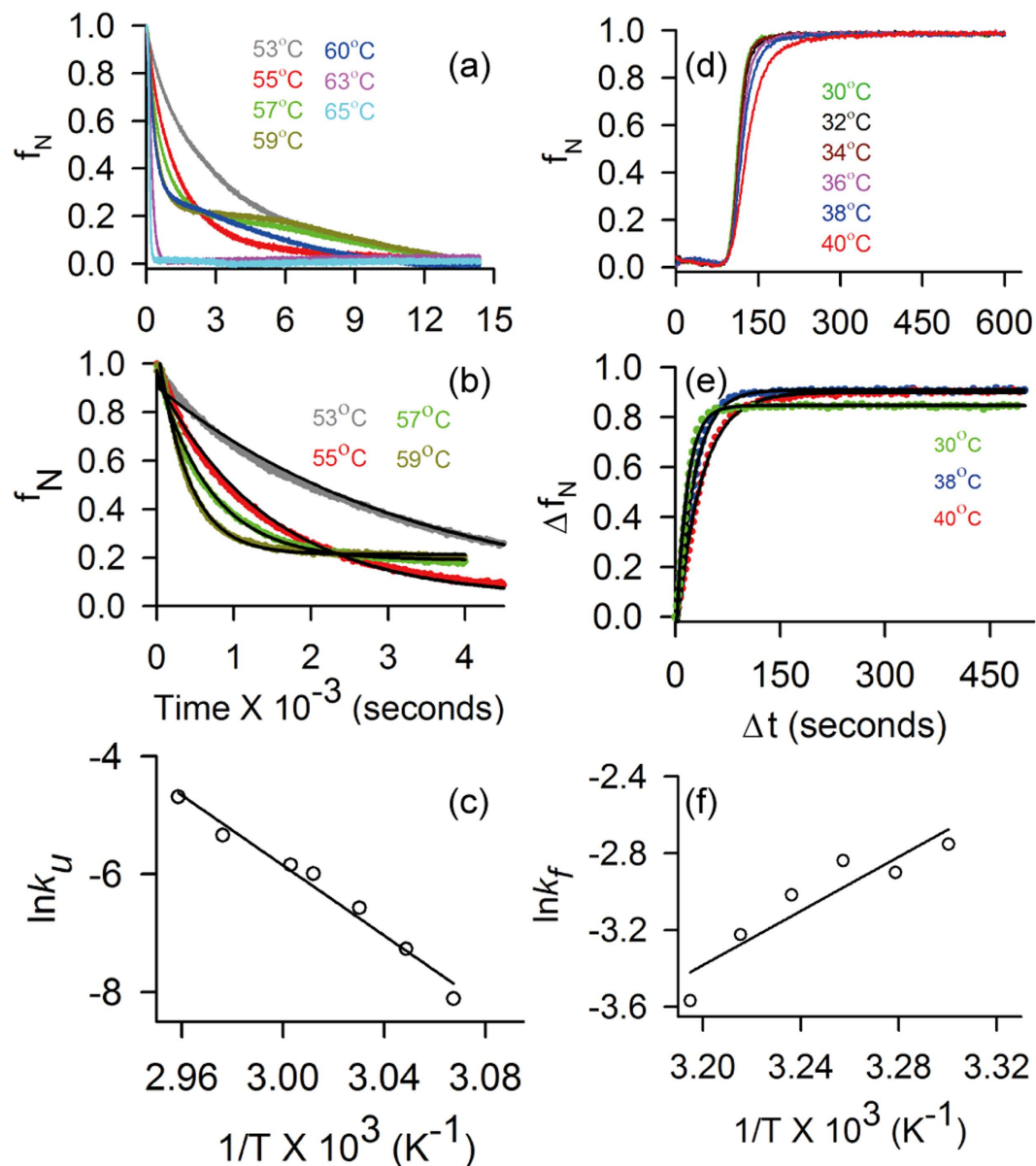
Figure 3. Kinetics of STY3178 unfolding and refolding. (**a**) Shows the folded protein fraction (f_N) as a function of time upon heating for temperatures 53 °C (grey), 55 °C (red), 57 °C (green), 59 °C (dark yellow), 60 °C (blue), 63 °C (magenta) and 65 °C (cyan). (**b**) Shows exponential fitting of f_N with time for temperatures 53 °C (grey), 55 °C (red), 57 °C (green) and 59 °C (dark yellow). (**c**) Arrhenius plot showing the linear dependence of $\ln k_u$ over 1/T (Kelvin^{-1}). (**d**) Shows the variation of f_N with time during protein refolding for temperatures 30 °C (green), 32 °C (black), 34 °C (brown), 36 °C (magenta), 38 °C (blue) and 40 °C (red). (**e**) The change in f_N ($\Delta f_N = f_N - f_R$, where f_R is the folded fraction from which the rise in f_N starts) versus time (Δt) at which f_N starts increasing, is fitted exponentially and shown for temperatures 30 °C (green), 38 °C (blue) and 40 °C (red). (**f**) The plot of $\ln k_f$ over 1/T (Kelvin^{-1}) shows linear dependence.

Molecular modeling. In the absence of a molecular structure, we propose a homology model of STY3178 to understand the stability of the protein, based on the template yfdX protein structure (PDB 3DZA) from *K. pneumoniae*. The sequence similarity of STY3178 with 3DZA protein is nearly 40%. The homology model contains residues A^{11}-Q^{183} excluding 21 N-terminal residues of the construct. At the C-terminal end of STY3178, there are 16 residues, for which no homology is obtained from the template. Thus, we add all these residues de-novo using Swiss-PdbViewer and minimize homology model for refinement. We perform MD simulation of this minimized homology model for 300 ns at 310 K. The MD simulated monomer is primarily helical containing total ten helices along with a two stranded antiparallel β-sheet as shown in Fig. 4a. The helices present in the structure are primarily forming two helix bundles where helices H2, H3, H7 and H8 form one bundle and helices H5, H6 and H10 forming the other.
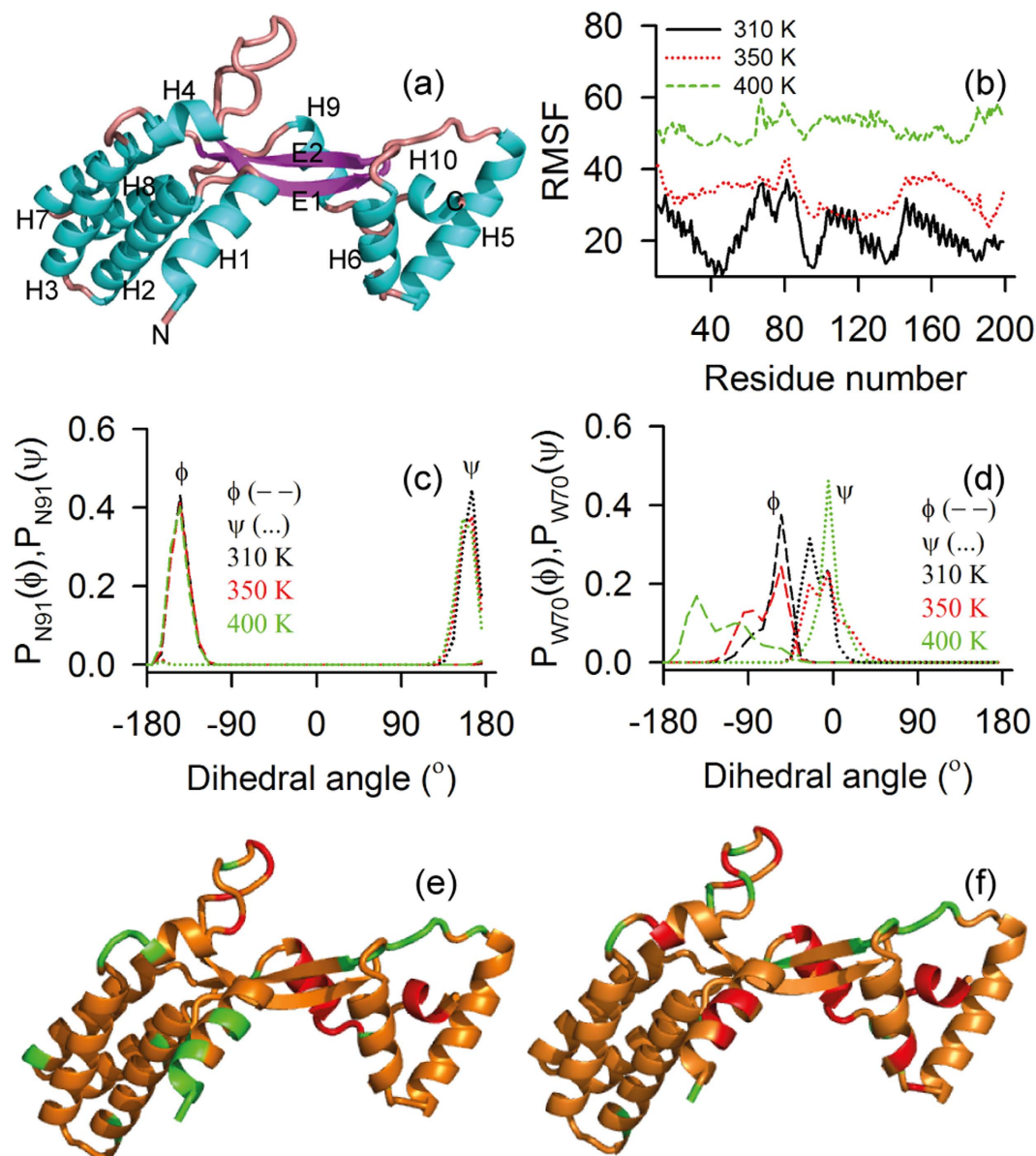
Figure 4. Molecular model and dihedral angle fluctuation of STY3178. (a) Cartoon representation of molecular model for STY3178. Ten α-helices (H1-H10, cyan) and two β-strands (E1 and E2, magenta) are marked along with the N- and C-terminals. **(b)** Root mean square fluctuation (RMSF) of each residue of the model in the equilibrated ensemble between 150–300 ns of simulated structure for temperatures 310 K (black), 350 K (red) and 400 K (green) temperatures. **(c,d)** Show the distribution of dihedral angles φ (dash) and ψ (dotted) at 310 K (black), 350 K (red) and 400 K (green) for residues N91and W70 respectively. **(e,f)** Show cartoon representation of the molecular model with residues as per the different degree of fluctuation at 350 K (r_{350}) and 400 K (r_{400}) respectively. Green indicates residues with r_{350}, $r_{400} \leq 1$, orange shows residues with r_{350} and r_{400} in the range 1–4 and the residues with r_{350}, $r_{400} > 4$ are shown in red.

Molecular dynamic simulations at elevated temperatures. Our experimental data show that this protein unfolds at elevated temperature without change in its assembly. This leads to test the thermal stability of the monomer model from MD simulations. We simulate structures at 310 K, 350 K and 400 K. We find that the root means square fluctuation (RMSF) increases with temperature (Fig. 4b). In particular there are two regions which show enhanced RMSF with temperature. The first region is part of H3 and H4 along with the adjacent loop residues and the other region comprises of E1, H7 and the adjacent loops. The backbone dihedral (φ and ψ) distribution for majority of residues do not show change with increase in temperature (Figs 4c and S2). Only a handful of residues show sensitivity in the φ and ψ distribution, which includes N14, D17, N18 from H1; D69, W70 and N71 from H4; A132 from H6; S182, Q183, S184 and V185 from H9; V193, H195, A197 and A198 of H10. An example of such change is shown for residue W70 in Fig. 4d and S184 in Supplementary Fig. S3. The Ramachandran plots for these residues indicate structural changes from helix to loop and β-sheet (Supplementary Fig. S4).
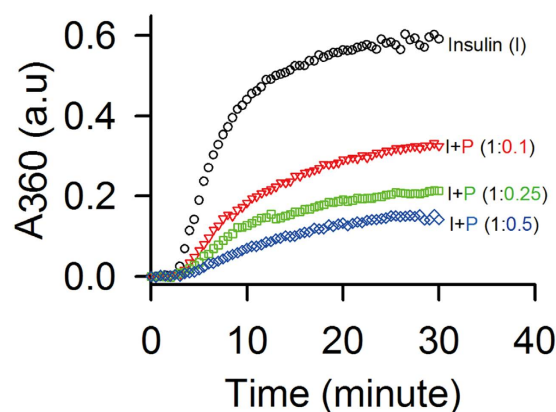
Figure 5. Chaperone activity assay. Absorbance at 360 nm (A_{360}) for DTT induced insulin (I) aggregation (black) is represented. Insulin aggregation in presence of STY3178 (P) at molar ratio 1:0.1 (red), 1:0.25 (green) and 1:0.5 (blue) are shown.

The relative standard deviations at 350 K and 400 K with respect to those at 310 K, given by $r_{350} = \sigma_{350}/\sigma_{310}$ and $r_{400} = \sigma_{400}/\sigma_{310}$ respectively, are shown in Fig. 4e,f for different residues, colour coded according to their values. Residues showing r_{350}, $r_{400} \leq 1$ for both ϕ and ψ are shown in green in Fig. 4e,f. Majority of residues belong to the region $1 < r_{350}$, $r_{400} < 4$ with moderate fluctuations in either ϕ or ψ or both at 350 K and 400 K as shown in orange in Fig. 4e,f. The residues, having r_{350}, $r_{400} > 4$ are marked in red which belong to the terminal helices H1, H9 and H10. Despite enhancement in fluctuations, the overall structure of STY3178 remains stable. To the contrary, we observe enhanced RMSF and loss of secondary structural element for lysozyme at 400 K from our simulation using the same force field as shown in Supplementary Fig. S5, in agreement with earlier report[35].

Discussion

We predict the function of STY3178 using bioinformatics tool MODexplorer[36]. The output results from MODexplorer server indicates nearly 50 proteins which show structural similarity >50% with STY3178. Among these 50 proteins most of the proteins have helices similar to STY3178. While some of the proteins with similar structures have no reported function, majority of them are either chaperones or assist chaperone activity. Our observations indicate that STY3178 may have chaperone activity. To verify such possibility, we monitor DTT induced aggregation of insulin B-chain in presence of STY3178. At 42 °C, we observe a steep rise in absorbance at 360 nm (A_{360}) with time followed by a saturation platue (Fig. 5) as reported earlier[37]. In presence of STY3178 at same temperature, we find lower absorbance at 360 nm with time, indicating less aggregation of insulin B-chain in presence of DTT. This prevention of aggregation is dependent on the net amount of STY3178 present in the reaction mixture. Figure 5 shows the plot of absorbance at 360 nm (A_{360}) versus reaction time (30 minutes) in absence (black) and presence of various STY3178 molar concentration ratios 0.1 (red), 0.25 (green) and 0.5 (blue) to insulin concentration. A systematic lowering of absorbance at 360 nm (A_{360}) with increasing concentration of STY3178 demonstrates its capability to prevent insulin aggregation. Thus, *in vitro* we indeed observe that STY3178 is capable of showing chaperone activity.

STY3178 sequence from Gene bank (gene ID gi|16758993:c3049965-3049366) is used in signalP 4.1 server[38] to identify any possible signal peptide sequence. The construct of STY3178 used in experiments do not contain 9 N-terminal residues. SignalP server predicts the first 12 residues in the N-terminal region of this protein as a signal peptide. Experimentally we observe that STY3178 is soluble in aqueous medium. This indicates that STY3178 cannot be a membrane embedded protein. Thus presence of predicted signal peptide in the N-terminus indicates a possibility of its localization in the periplasm. Exact cellular localization of this protein is not yet established. However, servers like Cello[39,40] and LocTree3[41] predicts its sub-cellular localization in the periplasm. There are many chaperones identified in the periplasm of bacteria like peptidyl-prolyl isomerases, disulphide bond isomerases etc[42–44]. Since ATP is absent in periplasm[42], the chaperones from periplasm are capable of performing their activity without ATP assistance. The predicted sub cellular localization as well as ATP independent chaperone activity of STY3178 both suggests its localization in the periplasm.

Conclusion

To summarize, our experimental studies reveal a reversible thermal unfolding and structural stability at elevated temperature for STY3178. The CD, steady state fluorescence and NMR data show that the protein is stable up to 45 °C and folds back even after heating at 80 °C. Further the time dependent CD measurements between 20 °C–70 °C show reversibility in unfolding with hysteresis. The kinetics of unfolding and refolding both follows the Arrhenius behavior with an activation energy barrier of ~246.9 kJ/mol and −58.66 kJ/mol, respectively. Our molecular dynamics simulation studies on the proposed model monomeric structure of the protein show stability at high temperature. Most importantly, we find an ATP independent chaperone activity capability of STY3178. This observation matches well with the bioinformatically predicted periplasmic localization of the protein. Our studies indicate that STY3178 may be an important protein through its chaperone activity for the bacterial cellular functions. Hence yfdX proteins need to be fully characterized to understand bacterial virulence.

Methods

Plasmid pET28a carrying the desired gene *sty3178* is overexpressed in *E. coli* using 0.2 mM isopropyl-β-D-thiogalactoside for 4 hours in a shaker (Innova 42 New Brunswick Scientific) as reported earlier[2]. Protein extraction and purification are done following the earlier reported protocol[2]. All the reported measurements are carried out in a buffer containing 50 mM phosphate (pH 7), 250 mM NaCl and 1 mM PMSF. In all the experiments, the protein concentration is calculated in terms of monomer molecular weight (MW = 23107.7 Da).

Circular Dichroism.   All the circular dichroism (CD) measurements are carried out in Jasco J-815 CD spectrometer equipped with peltier temperature control unit (Jasco). The sample concentration and path length for all the experiments are 10 μM and 3 mm, respectively. Every CD spectrum is acquired in the far UV-region (200–250 nm). Unfolding of the protein is monitored in the temperature range of 20 °C to 70 °C upon increasing temperature with increments of 10 °C and equilibration of 10 minutes at each (30 °C, 40 °C, 50 °C, 60 °C and 70 °C) temperature. Similar to unfolding experiments, refolding of the protein is achieved by decreasing temperature from 70 °C with decrements of 10 °C till 20 °C (60 °C, 50 °C, 40 °C, 30 °C, and 20 °C). Background correction is done by subtracting the ellipticity value of buffer in the far UV-region for each experiment at that temperature.

The transition from folded state to unfolded state is monitored by measuring the ellipticity values at 222 nm $[\theta_{222}]$ upon gradually heating from 20 °C to 70 °C. The fraction of the folded protein (f_N) at each temperature is calculated using the expression[29] $f_N = (\theta_D - \theta_T)/(\theta_D - \theta_N)$, where θ_N and θ_D are the measured ellipticity of the folded and the unfolded protein respectively. θ_T is the ellipticity at each temperature. Heating rates 0.5 °C/min, 1 °C/min, 2 °C/min, 3 °C/min, 4 °C/min and 5 °C/min are used in CD experiments for probing the transition from folded to unfolded state. Similarly we use six different cooling rates (0.5 °C/min, 1 °C/min, 2 °C/min, 3 °C/min, 4 °C/min and 5 °C/min) for refolding the protein.

Dynamic Light Scattering (DLS).   DLS data is acquired for the temperatures 20 °C, 30 °C, 40 °C, 50 °C and 60 °C using Nano-S Malvern instrument equipped with thermostat cell holder using methodology as described earlier[2]. Protein is equilibrated for 10 minutes at each temperature prior to acquiring data.

Unfolding Kinetics.   Thermal unfolding kinetics is followed from the change in ellipticity $[\theta_{222}]$ upon heating the protein for 4 hours at temperatures 53 °C, 55 °C, 57 °C, 59 °C, 60 °C, 63 °C and 65 °C. The buffer is pre-equilibrated at each temperature prior to mixing with protein. The dead time of the whole process is less than 15 sec. f_N is calculated at different times and the time dependence has been fitted using single exponential decay equation, $f_N = y_f + a e^{-k_u t}$, where k_u is rate of unfolding and t is time. The activation energy (E_a^u) of unfolding has been estimated using Arrhenius equation, $k_u \sim e^{-E_a^u/RT}$, where R is the universal gas constant and T is the absolute temperature.

Refolding Kinetics.   The kinetics of refolding is calculated from the ellipticity change $[\theta_{222}]$ of the unfolded protein at 222 nm. 10 μM protein is unfolded by heating at 70 °C for 10 minutes. The temperatures of the cell holder is set to six different temperatures (30 °C, 32 °C, 34 °C, 36 °C, 38 °C and 40 °C) allowing cooling of the unfolded protein and subsequent refolding. The time between the start of the experiment and rise of f_N is the lag time required for initiation of refolding. The lag time is primarily governed by the machine accuracy (99.5 sec-30 °C, 99.5 sec-32 °C, 99.5 sec-34 °C, 101 sec-36 °C, 99.5 sec-38 °C and 101 sec-40 °C). The entire set of data is collected for three repeated sets of experiments. We calculate the $\Delta f_N = f_N - f_R$, where f_R is the folded fraction from which the rise in f_N starts. The f_R values are typically small but non zero for different temperatures. Similarly, we define the rise time Δt by subtracting out the lag time. The plot of Δf_N versus Δt show single exponential rise, fitted with $f_N = y_0 + a(1 - e^{-k_f t})$ from which we estimate the rate of refolding (k_f). The activation energy of refolding (E_a^f) is obtained following Arrhenius equation by linear fitting of $\ln k_f$ with $1/T$.

Steady State Fluorescence.   Fluorescence experiments are performed in Jobin Yvon Horiba Fluorolog attached with a peltier unit (Wavelength Electronics). Sample concentration of 10 μM and slit width of 2 nm is used for excitation at 280 nm. Sample is heated gradually at different temperatures starting from 20 °C to 70 °C with increments of 5 °C and equilibration of 10 minutes at each temperature (25 °C, 30 °C, 35 °C, 40 °C, 45 °C, 50 °C, 55 °C, 60 °C, 65 °C and 70 °C). Similarly, the same sample is allowed to cool back gradually from 70 °C to the starting temperature (20 °C) to monitor its refolding. The final emission spectra are obtained by subtracting the corresponding buffer spectrum (blank) from that of the sample.

Nuclear Magnetic Resonance spectroscopy.   The uniformly ^{15}N-labeled protein sample is prepared by growing *E. coli* cells in M9 minimal media containing [15]NH$_4$Cl as the only source for nitrogen[2]. Extraction and purification protocols are same as done earlier[2]. NMR experiments are performed in a 600 MHz Varian spectrometer equipped with room temperature probe. [1]H-^{15}N heteronuclear single quantum coherence (HSQC) experiments are carried out for temperatures 25 °C, 30 °C, 35 °C, 40 °C and 45 °C.

Sample concentration used is 400 μM. The buffer condition used is 30 mM phosphate buffer (pH 7), 150 mM NaCl and 10% D$_2$O. All HSQC experiments are performed with 64 scans and 128 number of points. Data processing and analysis are done using NMRPIPE[45] and NMRVIEW[46].

Molecular model.   Homology model of the tertiary structure of STY3178 is generated using the servers Phyre2[47] and RaptorX[48–51]. Template used for the model generation is a yfdX protein from *K. pneumoniae* (PDB 3DZA) which is a homotetramer and contains four identical chains (A, B, C and D). Phyre2 used chain B and RaptorX used chain A of the template protein and generated identical models for STY3178. The homology model contains residues A[11]-Q[183] only. There is no homology obtained for residues S[184]-H[199]. We generate a model

for this region (S^{184}-H^{199}) using Swiss-PdbViewer[52] by adding one residue at a time and minimizing thereafter. Hydrogen atoms are added to the final model containing residues A^{11}-H^{199}. The model is solvated using explicit solvent model in a rectangular parallelepiped water box with dimension $69.4 \times 62.2 \times 73$ Å^3 and minimized using NAMD[53] for refinement after neutralizing with counter ions.

Molecular dynamics simulation of protein model. NAMD[53] is used for the molecular dynamics (MD) simulation of the minimized structure. CHARMM27[54] force field and TIP3P[55] water model is used for MD simulation performed at 1 atm pressure in isothermal-isobaric (NPT) ensemble following the standard protocol, with periodic boundary condition and 1 femtosecond (fs) time-step. Particle-mesh Ewald method is applied to deal with electrostatic interactions. Energy minimization of 10,000 steps is performed prior to 300 nanosecond (ns) simulation time with 29848 total number of atoms. The initial MD simulation is performed at 310 K temperature. An equilibrated structure from the 310 K ensemble is used for 350 K and 400 K temperature simulations. The root mean square deviation (RMSD) at all the temperatures is calculated with respect to the C_α-atoms over all the trajectories up to 300 ns. We then estimate the root means square fluctuation (RMSF) of C_α-atoms of each residue at temperatures 310 K, 350 K and 400 K over the equilibrated trajectories. The histogram distribution, Ramachandran plot and standard deviation (σ) of backbone dihedral angles ϕ and ψ of each residue is calculated over the trajectories 150–300 ns. The relative change in σ for each residue at 350 K ($r_{350} = \sigma_{350}/\sigma_{310}$) and 400 K ($r_{400} = \sigma_{400}/\sigma_{310}$) is then estimated with respect to that at 310 K.

MD simulation for the protein lysozyme (PDB 193L) is performed for 100 ns using the total number of 21002 atoms at 310 K and 400 K following the protocols mentioned above.

Chaperone activity assay. Thermal aggregation of 50 μM insulin in presence of 20 mM DTT is monitored from the absorbance at 360 nm (A_{360}). A 10 mm path length cell is used in BioSpectrometer (Eppendorf) for A_{360} measurement at 42 °C as described in earlier protocol[37]. The experiment is performed both in absence and in the presence of pure STY3178. Molar ratio of STY3178 used in the assay is 0.1:1, 0.25:1 and 0.5:1 with respect to insulin. The buffer used in the experiment is 50 mM phosphate (pH 7), 250 mM NaCl and 1 mM PMSF. Both the proteins are equilibrated at 42 °C for 10 minutes prior to addition of DTT.

References

1. Bateman, A., Coggill, P. & Finn, R. D. DUFs: families in search of function. *Acta Crystallogr, Sect F: Struct Biol Cryst Commun* **66,** 1148–1152 (2010).
2. Saha, P., Manna, C., Das, S. & Ghosh, M. Antibiotic binding of STY3178, a yfdX protein from Salmonella Typhi. *Nature Scientific Reports* **6,** 21305 (2016).
3. Nishino, K. & Yamaguchi, A. Overexpression of the response regulator evgA of the two-component signal transduction system modulates multidrug resistance conferred by multidrug resistance transporters. *J Bacteriol* **183,** 1455–1458 (2001).
4. Nishino, K., Inazumi, Y. & Yamaguchi, A. Global analysis of genes regulated by EvgA of the two-component regulatory system in Escherichia coli. *J Bacteriol* **185,** 2667–2672 (2003).
5. Ma, D. *et al.* Genes acrA and acrB encode a stress-induced efflux system of Escherichia coli. *Mol Microbiol* **16,** 45–55 (1995).
6. England, J. L. & Haran, G. Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back. *Annu Rev Phys Chem* **62,** 257–277 (2011).
7. Kishore, D., Kundu, S. & Kayastha, A. M. Thermal, chemical and pH induced denaturation of a multimeric beta-galactosidase reveals multiple unfolding pathways. *PLoS One* **7,** e50380 (2012).
8. Strucksberg, K. H., Rosenkranz, T. & Fitter, J. Reversible and irreversible unfolding of multi-domain proteins. *Biochim Biophys Acta* **1774,** 1591–1603 (2007).
9. Duy, C. & Fitter, J. Thermostability of irreversible unfolding alpha-amylases analyzed by unfolding kinetics. *J Biol Chem* **280,** 37360–37365 (2005).
10. Findlay, H. E., Rutherford, N. G., Henderson, P. J. & Booth, P. J. Unfolding free energy of a two-domain transmembrane sugar transport protein. *Proc Natl Acad Sci USA* **107,** 18451–18456 (2010).
11. Meersman, F. *et al.* Consistent picture of the reversible thermal unfolding of hen egg-white lysozyme from experiment and molecular dynamics. *Biophys J* **99,** 2255–2263 (2010).
12. Blumlein, A. & McManus, J. J. Reversible and non-reversible thermal denaturation of lysozyme with varying pH at low ionic strength. *Bba-Proteins Proteom* **1834,** 2064–2070 (2013).
13. Eisenberg, M. A. & Schwert, G. W. The reversible heat denaturation of chymotrypsinogen. *J Gen Physiol* **34,** 583–606 (1951).
14. Yu, M. H., Weissman, J. S. & Kim, P. S. Contribution of Individual Side-Chains to the Stability of Bpti Examined by Alanine-Scanning Mutagenesis. *Journal of Molecular Biology* **249,** 388–397 (1995).
15. Roychaudhuri, R., Sarath, G., Zeece, M. & Markwell, J. Reversible denaturation of the soybean Kunitz trypsin inhibitor. *Arch Biochem Biophys* **412,** 20–26 (2003).
16. Wu, Y. V. & Scheraga, H. A. Studies of soybean trypsin inhibitor. II. Conformational properties. *Biochemistry* **1,** 905–911 (1962).
17. Garcia, A. F., Garcia, W., Nonato, M. C. & Araujo, A. P. Structural stability and reversible unfolding of recombinant porcine S100A12. *Biophys Chem* **134,** 246–253 (2008).
18. Wallevik, K. Reversible denaturation of human serum albumin by pH, temperature, and guanidine hydrochloride followed by optical rotation. *J Biol Chem* **248,** 2650–2655 (1973).
19. Pico, G. A. Thermodynamic features of the thermal unfolding of human serum albumin. *Int J Biol Macromol* **20,** 63–73 (1997).
20. Benitez-Cardoza, C. G., Rojo-Dominguez, A. & Hernandez-Arana, A. Temperature-induced denaturation and renaturation of triosephosphate isomerase from Saccharomyces cerevisiae: evidence of dimerization coupled to refolding of the thermally unfolded protein. *Biochemistry* **40,** 9049–9058 (2001).
21. Chien, A., Edgar, D. B. & Trela, J. M. Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *J Bacteriol* **127,** 1550–1557 (1976).
22. Ishino, S. & Ishino, Y. DNA polymerases as useful reagents for biotechnology-the history of developmental research in the field. *Frontiers in microbiology* **5,** 465 (2014).
23. Shearstone, J. R. & Baneyx, F. Biochemical characterization of the small heat shock protein IbpB from Escherichia coli. *J Biol Chem* **274,** 9937–9945 (1999).
24. Basha, E., Jones, C., Wysocki, V. & Vierling, E. Mechanistic Differences between Two Conserved Classes of Small Heat Shock Proteins Found in the Plant Cytosol. *Journal of Biological Chemistry* **285,** 11489–11497 (2010).
25. Boudker, O., Todd, M. J. & Freire, E. The structural stability of the co-chaperonin GroES. *Journal of Molecular Biology* **272,** 770–779 (1997).

26. Schlapschy, M. et al. The periplasmic E. coli chaperone Skp is a trimer in solution: biophysical and preliminary crystallographic characterization. Biological chemistry 385, 137–143 (2004).
27. Scholz, C. et al. SlyD proteins from different species exhibit high prolyl isomerase and chaperone activities. Biochemistry 45, 20–33 (2006).
28. Panse, V. G., Swaminathan, C. P., Aloor, J. J., Surolia, A. & Varadarajan, R. Unfolding thermodynamics of the tetrameric chaperone, SecB. Biochemistry 39, 2362–2369 (2000).
29. Camarillo-Cadena, M., Garza-Ramos, G., Peimbert, M., Perez-Hernandez, G. & Zubillaga, R. A. Thermal denaturation of beta-glucosidase B from Paenibacillus polymyxa proceeds through a Lumry-Eyring mechanism. Protein J 30, 318–323 (2011).
30. Freire, E., van Osdol, W. W., Mayorga, O. L. & Sanchez-Ruiz, J. M. Calorimetrically determined dynamics of complex unfolding transitions in proteins. Annu Rev Biophys Biophys Chem 19, 159–188 (1990).
31. Tello-Solis, S. R. & Hernandez-Arana, A. Effect of irreversibility on the thermodynamic characterization of the thermal denaturation of Aspergillus saitoi acid proteinase. Biochem J 311, 969–974 (1995).
32. Kajitani, S., Fukuoka, M. & Sakai, N. Kinetics of thermal denaturation of protein in cured pork meat. Japan Journal of Food Engineering 12, 19–26 (2011).
33. Zoldak, G., Zubrik, A., Musatov, A., Stupak, M. & Sedlak, E. Irreversible thermal denaturation of glucose oxidase from Aspergillus niger is the transition to the denatured state with residual structure. Journal of Biological Chemistry 279, 47601–47609 (2004).
34. Banerjee, T. & Kishore, N. A differential scanning calorimetric study on the irreversible thermal unfolding of concanavalin A. Thermochim Acta 411, 195–201 (2004).
35. Eleftheriou, M., Germain, R. S., Royyuru, A. K. & Zhou, R. H. Thermal denaturing of mutant lysozyme with both the OPLSAA and the CHARMM force fields. Journal of the American Chemical Society 128, 13388–13395 (2006).
36. Kosinski, J., Barbato, A. & Tramontano, A. MODexplorer: an integrated tool for exploring protein sequence, structure and function relationships. Bioinformatics 29, 953–954 (2013).
37. Weinreb, O., van Boekel, M. A., Dovrat, A. & Bloemendal, H. Effect of UV-A light on the chaperone-like properties of young and old lens alpha-crystallin. Investigative ophthalmology & visual science 41, 191–198 (2000).
38. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature methods 8, 785–786 (2011).
39. Yu, C. S., Lin, C. J. & Hwang, J. K. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Sci 13, 1402–1406 (2004).
40. Yu, C. S., Chen, Y. C., Lu, C. H. & Hwang, J. K. Prediction of protein subcellular localization. Proteins 64, 643–651 (2006).
41. Goldberg, T. et al. LocTree3 prediction of localization. Nucleic Acids Res 42, W350–W355 (2014).
42. Wulfing, C. & Pluckthun, A. Protein folding in the periplasm of Escherichia coli. Mol Microbiol 12, 685–692 (1994).
43. Bodelon, G., Marin, E. & Fernandez, L. A. Role of periplasmic chaperones and BamA (YaeT/Omp85) in folding and secretion of intimin from enteropathogenic Escherichia coli strains. J Bacteriol 191, 5169–5179 (2009).
44. Sklar, J. G., Wu, T., Kahne, D. & Silhavy, T. J. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. Genes Dev 21, 2473–2484 (2007).
45. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277–293 (1995).
46. Johnson, B. A. & Blevins, R. A. NMR View: A computer program for the visualization and analysis of NMR data. J Biomol NMR 4, 603–614 (1994).
47. Kelley, L. A. & Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4, 363–371 (2009).
48. Peng, J. & Xu, J. B. RaptorX: Exploiting structure information for protein alignment by statistical inference. Proteins 79, 161–171 (2011).
49. Kallberg, M. et al. Template-based protein structure modeling using the RaptorX web server. Nat Protoc 7, 1511–1522 (2012).
50. Ma, J. Z., Wang, S., Zhao, F. & Xu, J. B. Protein threading using context-specific alignment potential. Bioinformatics 29, 257–265 (2013).
51. Peng, J. & Xu, J. A multiple-template approach to protein threading. Proteins 79, 1930–1939 (2011).
52. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714–2723 (1997).
53. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. J Comput Chem 26, 1781–1802 (2005).
54. Brooks, B. R. et al. CHARMM: the biomolecular simulation program. J Comput Chem 30, 1545–1614 (2009).
55. Mahoney, M. W. & Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. Journal of Chemical Physics 112, 8910–8922 (2000).

Acknowledgements

Author Contributions

M.G. and J.C. planned the research. P.S. and C.M. carried out the experiments. P.S. performed the simulations. J.C. provided the support for computational work. M.G., J.C. and P.S. analyzed the results and wrote the paper. All authors reviewed the manuscript.

Additional Information

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

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

How to cite this article: Saha, P. et al. Reversible thermal unfolding of a yfdX protein with chaperone-like activity. Sci. Rep. 6, 29541; doi: 10.1038/srep29541 (2016).