

# SCIENTIFIC REPORTS



OPEN

## Thermal stabilization of the deglycating enzyme Amadoriase I by rational design

Federica Rigoldi<sup>1</sup>, Stefano Donini<sup>2</sup>, Francesca Giacomina<sup>3</sup>, Federico Sorana<sup>3</sup>, Alberto Redaelli<sup>1</sup>, Tiziano Bandiera<sup>3</sup>, Emilio Parisini<sup>2</sup> & Alfonso Gautieri<sup>1</sup>

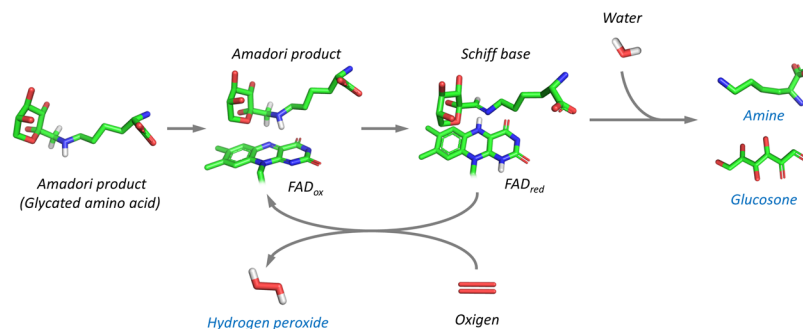
Amadoriases are a class of FAD-dependent enzymes that are found in fungi, yeast and bacteria and that are able to hydrolyze glycated amino acids, cleaving the sugar moiety from the amino acidic portion. So far, engineered Amadoriases have mostly found practical application in the measurement of the concentration of glycated albumin in blood samples. However, these engineered forms of Amadoriases show relatively low absolute activity and stability levels, which affect their conditions of use. Therefore, enzyme stabilization is desirable prior to function-altering molecular engineering. In this work, we describe a rational design strategy based on a computational screening method to evaluate a library of potentially stabilizing disulfide bonds. Our approach allowed the identification of two thermostable Amadoriase I mutants (SS03 and SS17) featuring a significantly higher  $T_{50}$  (55.3 °C and 60.6 °C, respectively) compared to the wild-type enzyme (52.4 °C). Moreover, SS17 shows clear hyperstabilization, with residual activity up to 95 °C, whereas the wild-type enzyme is fully inactive at 55 °C. Our computational screening method can therefore be considered as a promising approach to expedite the design of thermostable enzymes.

Amadoriases, also known as Fructosyl Amino Acid Oxidases (FAOX)<sup>1–3</sup>, are a class of enzymes that are found in fungi and bacteria and that are able to cleave low molecular weight Amadori product (i.e., glycated amino acids) to yield a free amine, glucosone and hydrogen peroxide<sup>4,5</sup> (Fig. 1).

Today, members of this class of enzymes, commonly referred to as Fructosyl Peptide Oxidases (FPOX)<sup>6–8</sup>, are used in the detection of glycated haemoglobin (HbA1c), a long term marker for diabetes (the half-life of HbA1c is 3 months). Such diabetes monitoring kits are based on a first proteolytic digestion of HbA1c, which releases single amino acids from the protein, including its N-terminal glycation-prone valine. Then, the FPOX enzyme binds glycated valine and hydrolyzes it producing hydrogen peroxide which, in turn, is measured in a colorimetric assay using horseradish peroxidase and a suitable chromophore<sup>9</sup>. Using a similar mechanism, the Amadoriase I enzyme is of potential use in the detection of glycated albumin, a short to mid-term glycemic marker for diabetes (glycated albumin has a 3 weeks half-life)<sup>10</sup>. However, a common problem with biosensors is the long-term stability of their biological components. Hence, a stabilized Amadoriase enzyme will improve biosensor stability during transport and storage and the overall shelf-life.

Interestingly, Amadoriase enzymes are also regarded as promising therapeutic tools for the prevention or the reduction of protein glycation in biological tissues<sup>11</sup>. Glycation is the spontaneous, non-enzymatic and irreversible reaction between a sugar moiety and a protein, leading to a covalent adduct<sup>12</sup>. By modifying the chemistry of functional proteins<sup>13,14</sup>, glycation leads to a cascade of adverse clinical outcomes including arterial stiffening<sup>15</sup>, atherosclerosis<sup>16</sup>, nephropathy<sup>17</sup>, retinopathy<sup>18</sup> and neuropathy<sup>19</sup>. So far, the use of Amadoriases to prevent protein glycation is precluded because these enzymes show no significant activity on intact proteins<sup>20,21</sup> owing to their buried active site location and to the narrow tunnel that provides access to the catalytic pocket<sup>18,19</sup>. A stabilized form of Amadoriase may therefore help to set the stage for a further, more extensive engineering approach to the design of protein deglycation tools.

<sup>1</sup>Biomolecular Engineering Lab, Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano, Italy. <sup>2</sup>Center for Nano Science and Technology @Polimi, Istituto Italiano di Tecnologia, Via G. Pascoli 70/3, 20133, Milano, Italy. <sup>3</sup>D3-PharmaChemistry, Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genova, Italy. Federica Rigoldi and Stefano Donini contributed equally to this work. Correspondence and requests for materials should be addressed to E.P. (email: [emilio.parisini@iit.it](mailto:emilio.parisini@iit.it)) or A.G. (email: [alfonso.gautieri@polimi.it](mailto:alfonso.gautieri@polimi.it))



**Figure 1.** Schematic of the amino acid deglycation reaction catalyzed by Amadoriase I. The enzyme catalyzes the oxidation of the C-N bond between the nitrogen of the amino acid moiety of the Amadori product and C1 of the fructosyl portion. The reaction results in the Schiff base which is hydrolyzed to yield glucosone and a free amino acid. The reduced FAD is oxidized by an oxygen molecule with the release of hydrogen peroxide. Only relevant hydrogen atoms are shown. Blue font highlights the products of the enzymatic catalysis.

A further potential application of thermostable Amadoriases concerns the prevention of acrylamide formation in processed food. Thermal treatments used in food manufacturing (such as baking, toasting, frying, and roasting) boost the Maillard reaction between reducing sugars and amino acids, which gives aromas and taste. However, the reaction between sugars and asparagine amino acids yields acrylamide, a carcinogenic compound<sup>22,23</sup>. This outcome is particularly relevant in fried potatoes, bakery products and coffee, and currently there are no viable strategies to mitigate acrylamide formation while preserving the desired properties of processed food<sup>24,25</sup>. A thermostable Amadoriase enzyme is a potential tool to be used in food processing to limit the Maillard reaction on single amino acids and thus acrylamide formation.

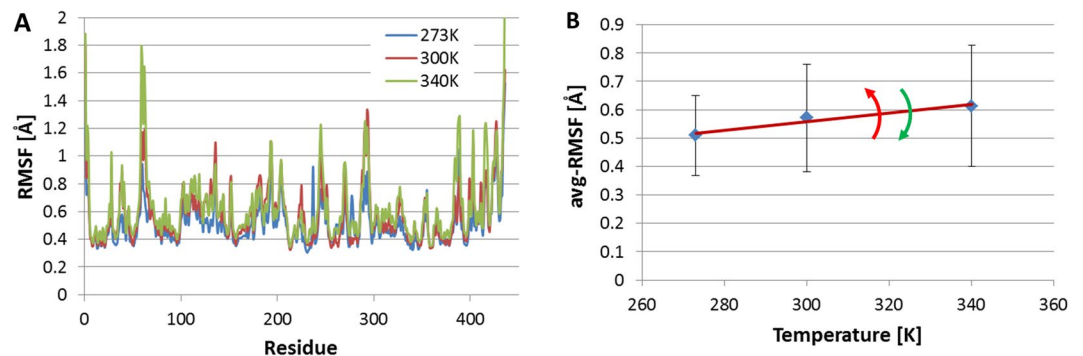
All the aforementioned applications of Amadoriases can benefit from or necessitate the stabilization of the enzymes. In this study, we apply a novel rational design approach to the stabilization of wild-type Amadoriase I from *Aspergillus fumigatus*. In general, any modification of the properties of a protein can in principle be achieved through the use of different rational approaches. Indeed, several computational methods to predict protein stability or changes in stability upon natural sequence mutation have been developed and tested<sup>26,27</sup>. However, the performance of these popular methods is still relatively weak<sup>28</sup>. Usually, the rational design of proteins with improved stability involves optimization of charge-charge interactions, saturation mutagenesis of residues with high crystallographic B-factors<sup>29</sup>, methods based on protein simulation and calculation of free energies<sup>30</sup>, phylogenetic analysis<sup>31</sup>, comparison to homologous proteins including the ultra-stable proteins of thermophilic organisms<sup>32</sup>, or structure-based computational design<sup>33,34</sup>. As expected, all these methods have a relatively high probability of inadvertently introducing disruptive mutations<sup>33,35</sup>. Hence, efforts to stabilize large proteins have also been based on library approaches aimed at identifying optimal combinations of stabilizing mutations<sup>36–38</sup>. However, such approaches are laborious and impractical for proteins without established medium-to-high-throughput screening protocols.

In an effort to address this issue efficiently, we developed a high-throughput computational screening method based on Molecular Dynamics (MD) simulations whereby the evaluation of a library of potentially stabilizing disulfide bonds (obtained by SSBOND software<sup>39</sup>) allows for the selection of a few candidates to be produced and tested experimentally. Here, we report on the identification, production and enzymatic characterization of four Amadoriase I mutants, two of which show a remarkable increase in thermal stability compared to the wild type enzyme.

## Results

**Computational design and screening of disulfide bonds.** Using the software SSBOND<sup>39,40</sup> we obtained a list of 19 possible disulfide bonds sites. Starting from the crystal structure of Amadoriase I (PDB id: 4WCT)<sup>41</sup>, we built the molecular models of the wild-type (WT) and those of the 19 Amadoriase I variants (SS01 to SS19), each featuring a different disulfide bond, and we screened them using MD simulations. The root mean square fluctuation (RMSF) was calculated for three different temperatures (270, 300 and 340 K) (Fig. 2 Panel A). The average RMSF (avg-RMSF) as a function of the temperature (Fig. 2 Panel B) was then interpolated with a linear equation. The slope ( $\lambda$ ) represents the index that we used to discriminate the stabilized mutants. SS-variants with a larger  $\lambda$  with respect to WT were discarded while those with  $\lambda$  lower than that of the WT were selected for experimental production.

As shown in Table 1, 13 out of the 19 possible disulfide bonds were excluded on the basis of a  $\lambda$  value larger than  $\lambda_{WT}$ . One variant (SS10) was excluded due to the poor fitting of avg-RMSF vs. temperature. We also excluded SS08 because the disulfide involved only residues from the rather floppy C-terminus and thus unlikely to stabilize the whole structure. At the end of our MD-based screening, four mutants (SS03, SS07, SS11 and SS17) were selected for experimental characterization (Fig. 3). The MD-based evaluation of each enzyme variant took approximately 10 h on a NVIDIA Tesla K80. We note that MD simulations of few nanoseconds are enough to capture local conformational changes and small loop rearrangements, while larger domain relaxation would require millisecond-long simulations, which are beyond the current state-of-the-art.



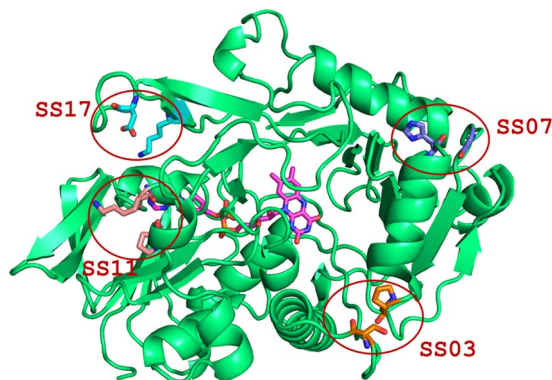
**Figure 2.** MD-based screening of mutants. **(A)** RMSF of WT-Amadoriase I at different temperatures, showing how higher temperatures induce an increase in the baseline but also a steep increase in specific regions of the protein (e.g., residue from 60 to 70). **(B)** avg-RMSF of WT-Amadoriase I at increasing temperatures. Here, the slope is considered as a proxy of the sensitivity to thermal stress and it is used to filter out mutations that badly affect thermal stability. In particular, enzyme mutants showing larger slopes are considered less stable than the WT, whereas enzyme mutants showing smaller slopes are considered as promising thermostable variants and thus experimentally tested.

Enzyme	Mutated residues	$\lambda$	$R^2$	p-value
WT	—	0.00151	0.953	
SS01*	S10, H216	0.00162	0.965	0.0407
SS02*	A29, L179	0.00201	0.996	0.0013
SS03	S67, P121	0.00151	0.995	0.8234
SS04*	A74, I399	0.00242	0.997	0.0004
SS05	Y95, A168	0.00187	0.993	0.8326
SS06*	G100, P260	0.00202	0.960	0.0028
SS07*	H106, G150	0.00115	0.997	0.0022
SS08*	K438, E441	0.00138	0.984	0.0187
SS09*	F148, W151	0.00181	0.972	0.0056
SS10*	V202, L263	0.00063	0.426	0.0021
SS11*	K233, P404	0.00134	0.924	0.0202
SS12*	K234, A407	0.00183	0.999	0.0027
SS13*	Q247, S337	0.00173	0.990	0.0070
SS14*	M248, V253	0.00202	0.939	0.0036
SS15*	C290, S304	0.00221	0.926	0.0026
SS16*	G432, Q435	0.00166	0.888	0.0475
SS17*	D295, K303	0.00118	0.993	0.0027
SS18*	R298, Q301	0.00187	0.999	0.0022
SS19*	L324, A332	0.00170	0.997	0.0080

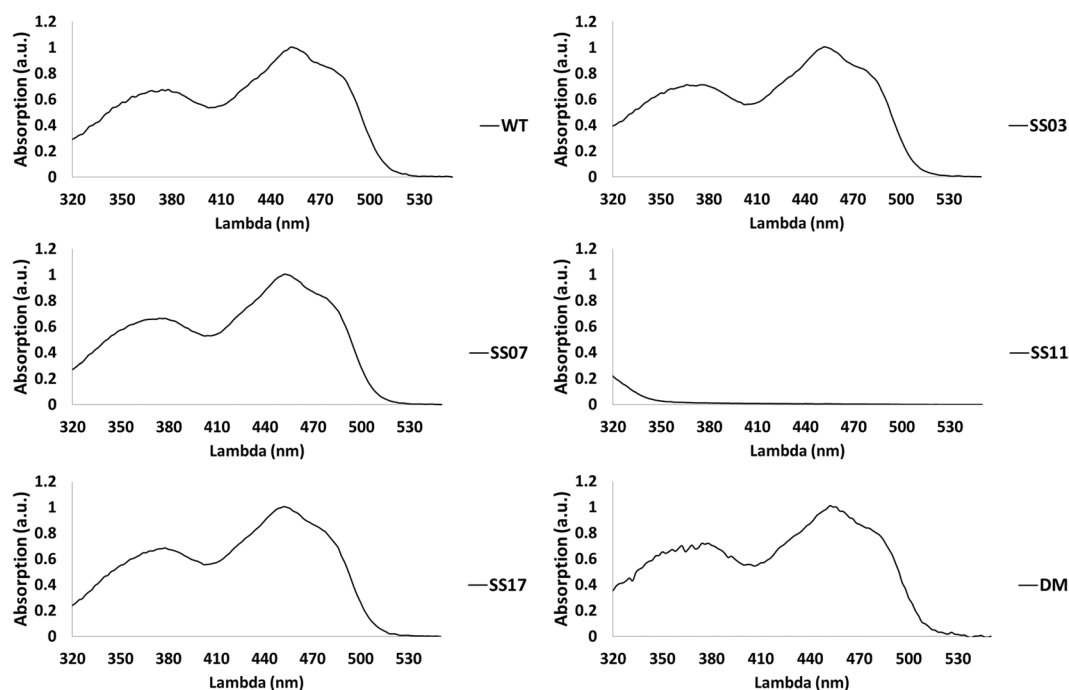
**Table 1.** Screening of SS-variants. As a result of the comparison of the  $\lambda$  value for the 19 SS-variants with  $\lambda_{WT}$ , four mutants (SS03, SS07, SS08, SS17) were selected for experimental production and characterization. Statistically significant differences between the  $\lambda$  value of SS-variant and  $\lambda$  of WT are marked with (\*).

**Production, purification and biochemical characterization of the Amadoriase variants.** Based on our computational design and screening method, we proceeded with the production and characterization of the four most promising thermostable mutants (SS03, SS07, SS11, and SS17). In addition, we produced and characterized a mutant species (SSDM) featuring the combination of the two most effective disulfide bonds resulting from the first round of experimental characterization (i.e., SS03 and SS17).

We obtained pure and active Amadoriase I SS03, SS07, SS17, and SSDM mutants with an average yield of 30 mg per liter of bacterial culture. In contrast, the SS11 variant showed much lower expression levels (less than 1 mg/L) and no detectable activity. The four highly expressed proteins exhibit a bright yellow color. Their absorption spectrum, which features two major absorbance peaks, one at 368 and one at 454 nm (Fig. 4), indicates the presence of flavin as prosthetic group. Hence, the absence of these signature peaks in the absorption spectrum of the SS11 variant suggests no binding of FAD moiety by the protein. This feature, combined with the dramatically reduced amount of soluble protein expressed and the complete enzyme inactivation, suggests that the cysteine residues introduced in position 233 and 404 highly destabilize the enzyme. For these reasons, we discarded the SS11 mutant and we did not proceed any further with its characterization.



**Figure 3.** Amadoriase I enzyme and selected mutations. The crystal structure of WT Amadoriase I (PDB code: 4WCT) is shown in cartoon representation, while the residues that were mutated to Cys in the selected SS-variants are represented in sticks (for SS03 residues S67 and P121, for SS07 residues H106 and G150, for SS11 residues K233 and P404, for SS17 residues D295 and K303). The FAD cofactor is shown in purple sticks.



**Figure 4.** UV-vis absorption spectra. The UV-visible absorption spectra of WT and SS03, SS07, SS17, and SS19 variants show two prominent peaks (at 368 nm and 454 nm) typical of bound FAD cofactor. Conversely, the SS11 does not show the two relevant regions of the UV-visible spectrum, indicating that it does not bind the flavin cofactor.

Apparent steady-state parameters ( $K_m$ ,  $k_{cat}$  and  $k_{cat}/K_m$ ) calculated for all the mutants towards fructosyl-lysine were consistent with those reported in the literature for wild-type Amadoriases<sup>9,20,42</sup> (Table 2). For the SS03 and the SS17 species, the mutations did not significantly affect the apparent steady state parameters relative to the WT. On the other hand, the SS07 mutant exhibited a two-fold increase in  $K_m$  as well as a four-fold decrease in catalytic efficiency compared to the WT. This may be due to a modification of the access tunnel to the catalytic pocket of the enzyme as a result of the introduction of a disulfide bond between residues H106 and G150, which are in close proximity to the entrance of the tunnel.

To support this hypothesis, we compared substrate tunnel dimension between the wild type enzyme and the SS07 mutant using CAVER 3.0.2 software<sup>43</sup>. For the analysis of the geometry of the tunnel, we removed the ions and water molecules that were included in the MD runs. In particular, we applied the same protocol developed in one of our previously published works<sup>44</sup>, which allowed us to identify the tunnel as a mean of the calculated tunnels in each frame of the MD trajectory. As a starting point for tunnel calculation we used the coordinate of the N5 atom of flavin, the enzyme's cofactor. As a probe, we defined a sphere of 1.4 nm radius, a value comparable to that of the cyclic form of glucose. To describe the surface of the protein correctly, we used a shell depth of 6 Å

Enzyme	$K_m$ [mM]	$k_{cat}$ [ $s^{-1}$ ]	$k_{cat}/K_m$ [ $s^{-1} mM^{-1}$ ]
WT	$0.51 \pm 0.19$	$21.55 \pm 3.08$	$41.68 \pm 16.48$
SS03	$0.34 \pm 0.13$	$21.90 \pm 2.73$	$64.17 \pm 25.5$
SS07	$0.99 \pm 0.22$	$11.97 \pm 1.69$	$12.02 \pm 2.86$
SS17	$0.68 \pm 0.18$	$22.91 \pm 2.12$	$33.34 \pm 10.16$
SSDM	$0.97 \pm 0.22$	$14.50 \pm 1.67$	$14.85 \pm 3.85$

**Table 2.** Comparison of the apparent steady state parameters for wild type Amadoriase I to the mutated enzymes using fructosyl-lysine as substrate. Experiments were performed at 25 °C, as described in the material and methods section.

Enzyme	Average bottleneck [ $\text{\AA}$ ]
WT	$1.83 \pm 0.17$
SS07	$1.57 \pm 0.20$

**Table 3.** Comparison of the size of the substrate tunnel bottleneck in the wildtype enzyme and in the SS07 mutant.

and a shell radius of 4  $\text{\AA}$ . We set the tunnel clustering threshold to 5  $\text{\AA}$ . Tunnel calculations were done at 5 ns of the MD simulations at 300 K. Default values were used for all the other parameters. For comparison, Table 3 reports the dimensions of the substrate tunnel bottleneck of the wild type enzyme and the SS07 mutant.

Our results show that MD simulations predict a local structural modification between two enzymes in the region of the substrate tunnel (see Fig. S1 in Supplementary Information). We suppose that this change could affect both the catalytic efficiency (since it involves the ligand pathway to the active pocket) and the thermal stability of the enzyme.

It is worth noting that, similarly to the SS07 mutant, also SSDM shows decreased catalytic efficiency and substrate affinity, suggesting that either the presence of both disulfide bonds results in some local structural perturbation that affects enzymatic activity or that no actual double SS bond formation occurs and the enzyme becomes partially destabilized by the introduced mutations.

The temperature effect on the WT and the mutants' activity is shown in Fig. 5. For all enzyme variants we report the  $T_{50}$ , i.e. the temperature at which the enzymes lose 50% of the activity with respect to the activity at 25 °C (Table 4). With the exception of the SS07 mutant, which shows a decreased stability, mutants SS03 and SS17 display a significant increase in  $T_{50}$  compared to the WT enzyme (3 °C and 8 °C, respectively). Interestingly, while all the enzymes, including the WT, lose completely their activity at temperatures  $\geq 60$  °C, the SS17 mutant retains a 50% residual activity at 60 °C and shows residual activity up to 95 °C. To confirm disulfide bond formation, we performed the same experiments also in the presence of 100 mM 1,4-dithiothreitol (DTT). Indeed, the reduced forms of all the SS-mutants lose their improved thermal resistance and behave very similarly to the WT. On the other hand, the SS07 is more stable in a reducing environment, further confirming the detrimental effect of the introduced disulfide bond in this species. Interestingly, the SSDM, which potentially carries both disulfide bonds of SS03 and SS17 does not display any cooperative effect relatively to the two single mutants; in fact, the measured  $T_{50}$  is consistent with the one obtained for the SS03 mutant, suggesting that only the C67-C121 disulfide bond may be actually formed.

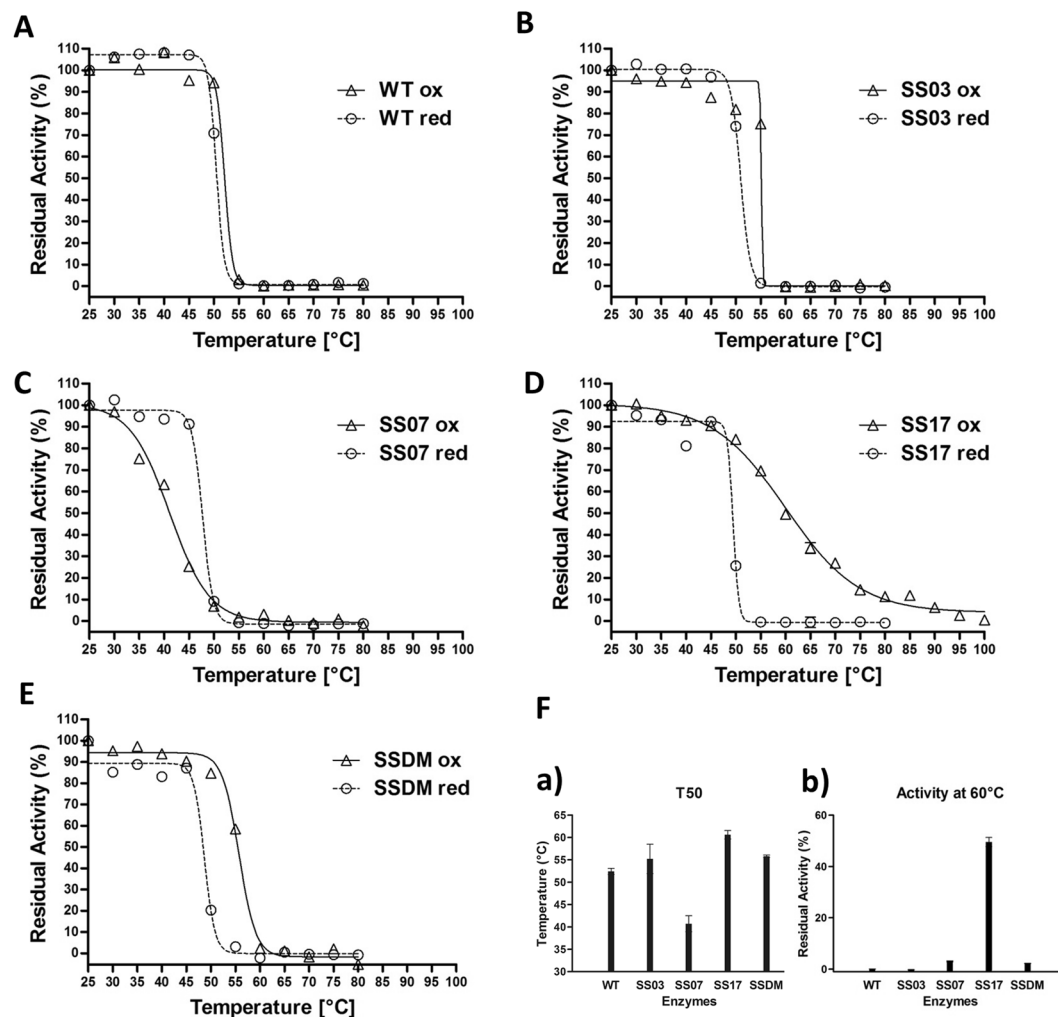
The pH-activity profiles for the WT and all the thermally stabilized enzymes described herein (SS03, SS17, and SSDM) feature a rather similar behavior, showing a decrease in activity below pH 8.0 and a plateau above this value (Fig. 6). Therefore, the increase in thermal stability due to the introduction of disulfide bonds is not paralleled by any increase in resistance to a wider range of pH values compared to the WT.

**Crystal structures.** We used X-ray crystallography to confirm the introduction of the designed disulfide bonds in the mutant enzymes. The crystal structures of the SS03 and the SS17 mutants were determined at 2.15 and 2.85  $\text{\AA}$  resolution, respectively. Both structures feature two independent molecules in the asymmetric unit. However, while the SS03 mutant crystallizes in the same space group (orthorhombic  $P2_12_12_1$ ) as the wild type enzyme (PDB code: 4WCT), the SS17 variant shows a different packing arrangement as it crystallizes in the monoclinic space group  $P2_1$  (Fig. S2 of Supplementary Information). The disulfide bonds as they appear in the electron density maps of our two mutants are shown in Fig. 7.

The overall fold of the two mutants is very similar to that of the wild-type enzyme, with an RMSD for main chain atoms of 0.268 and 0.804  $\text{\AA}$  for SS03 and SS17, respectively.

No significant variation in the geometry of the catalytic pocket can be observed between the two mutants and the wild type enzyme, demonstrating that indeed the introduction of these disulfide bonds, which is intended for thermal stabilization only, does not alter the architecture of the active site of the enzyme (Fig. 8, Panel A and B). This observed structural invariability in the catalytic pocket correlates with the activity profile of the mutant enzymes, which are by and large similar to the wild type species up to the temperature at which the natural Amadoriase I maintains its fold and activity.

To assess the results of the MD simulations in terms of structural conformation, we calculated the Root Mean Square Deviation (RMSD) for main-chain atoms between the crystallographic structures and the final conformation obtained by MD at 300 K after a 5 ns-long simulation using the ALIGN function implemented in Pymol<sup>45</sup>. We



**Figure 5.** Temperature effect on the activity of the WT Amadoriase I and the mutants. (A–E) Residual activity of the oxidized form (triangle) and the reduced form (circle) as a function of the temperature for the WT enzyme and the tested SS-mutants. Enzymes were treated at temperatures ranging from 25 °C to 100 °C for 10 min before the activity assay. The residual activity at 25 °C, for each enzyme, is assumed as 100%. (F) Panel (a) Temperature at which each enzyme loses half of its initial activity (measured at 25 °C) Panel (b) Residual activity at 60 °C. The reported data are the mean values of three replicates  $\pm$ SD (standard deviation).

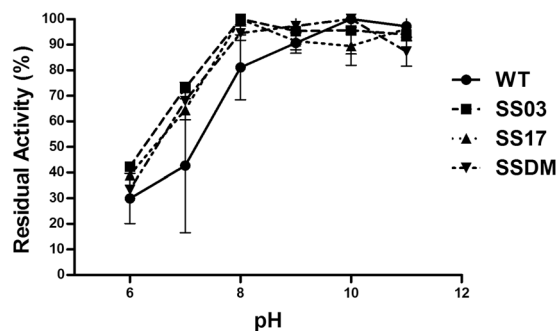
Enzyme	$T_{50}$ [°C]	$\Delta T_{50}$ [°C]
WT	52.40 $\pm$ 0.69	—
SS03	55.25 $\pm$ 3.28	+2.85
SS07	40.76 $\pm$ 1.80	-11.64
SS17	60.62 $\pm$ 0.95	+8.22
SSDM	55.79 $\pm$ 0.25	+3.39

**Table 4.** Absolute  $T_{50}$  and difference with respect to WT Amadoriase ( $\Delta T_{50}$ ). The reported data are the mean values of three replicates  $\pm$ SD (standard deviation).

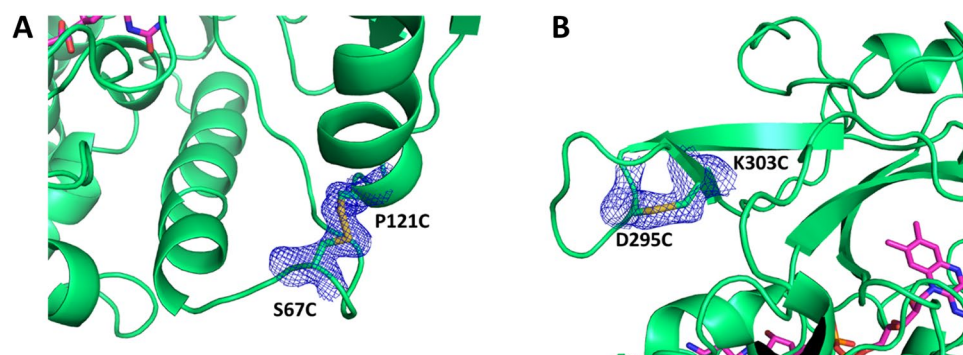
also performed the same analysis for the wild type enzyme (PDB code: 4WCT). The RMSD data are summarized in Table 5.

Our results show that computational conformations are in very good agreement with the experimental structures. This demonstrates that our MD simulations are reliable and that the global conformations of the two mutants species are very similar to the native form of Amadoriase I.

Focusing on the region where the mutations were introduced, the SS03 mutant (S67C/P121C) is highly superimposable to the wild type enzyme (residues 62 to 72 and 116 to 126, RMSD = 0.842) while the SS17 variant (D295C/K303C) shows a structural rearrangement in the region from residue 290 to 308, with an RMSD of 1.657 Å (Fig. 8, Panel C and D).



**Figure 6.** Effect of pH on the activity of the WT Amadoriase I and the thermally stabilized mutants. Activity was measured at 25 °C in a pH range from 6 to 11. Experimental conditions are described in the materials and methods section. The activity is normalized to the maximum activity for each enzyme. The reported data are the mean values of three replicates  $\pm$ SD (standard deviation).



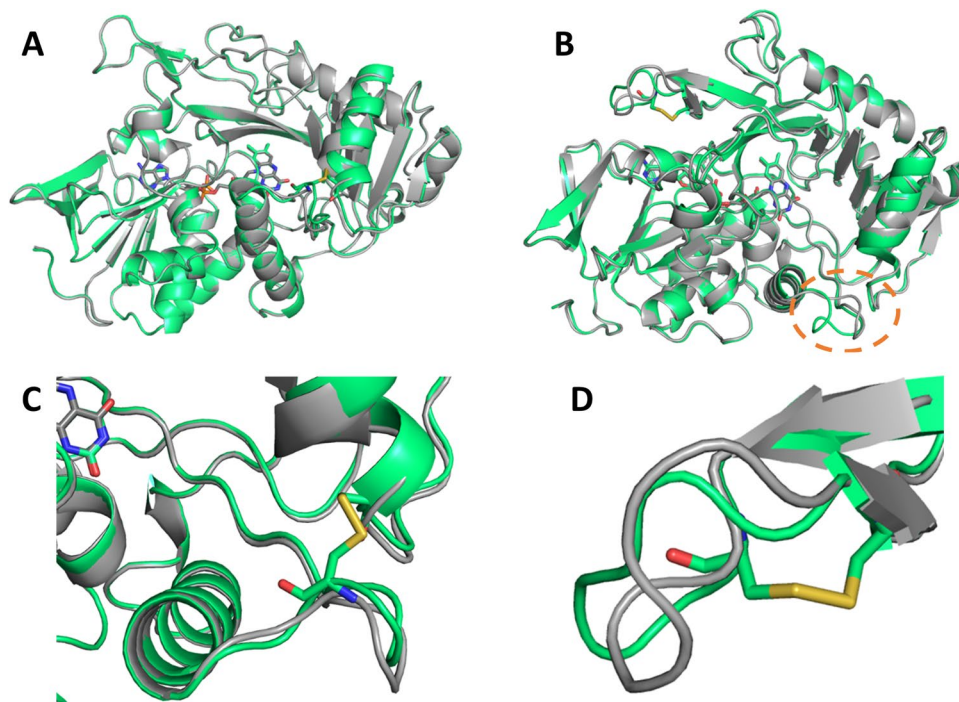
**Figure 7.** Detailed view of the electron density map in the region of the double cysteine mutation for the SS03 enzyme variant (Panel A) and for the SS17 enzyme variant (Panel B). Both maps show the clear formation of the disulfide bonds.

Interestingly, while the introduction of the disulfide bond in the SS03 mutant does not cause any significant structural rearrangement in any portion of the enzyme, the formation of the disulfide bond in the SS17 mutant affects the conformation of one of the loops that define the boundaries of the tunnel leading to the catalytic pocket (residues 63 to 69) (Fig. 8, Panel B, orange circle). The RMSD between the SS17 mutant and the wild type enzyme for this short region is 2.745 Å. It is worth noting that despite this displacement, this portion of Amadoriase I is still structurally more stable than the corresponding loop in the crystal structure of Amadoriase II (PDB code: 3DJJ), which could not be determined due to its high flexibility<sup>9</sup>.

## Discussion

Enhanced thermal stability is generally an essential requirement for those enzymes that are used as biocatalysts in industrial processes occurring at high temperatures. To this end, an enzyme can be modified by enhancing glycosylation<sup>46,47</sup>, or by introducing disulfide bridges<sup>48</sup>, salt-bridges<sup>49</sup>, or hydrogen bonds<sup>50</sup>. In particular, the introduction of disulfide bonds provide considerable stability to proteins by locking their fold in a well-defined local or global conformation<sup>51,52</sup>. Improved enzyme thermostability can usually be achieved either by structure-guided rational approaches<sup>53</sup> or through a directed evolution campaign<sup>54,55</sup>. However, both these methods can be extremely time-consuming and highly labor-intensive. To the best of our knowledge, the only prior attempt to stabilize an Amadoriase enzyme has been carried out using a directed evolution approach that provided a  $T_{50}$  increase of  $\approx 5$  °C as a result of the combination of five point mutations<sup>56</sup>.

Here, with the goal of providing thermal stabilization to the Amadoriase I enzyme without incurring in an extensive experimental validation phase, we introduced a computational high-throughput screening method based on MD simulations, which is used to evaluate a library of mutations. Our library of potentially stabilizing disulfides was obtained with SSBOND code<sup>39,40</sup>. Based on the backbone and sidechain geometries of the WT, the software determined the residue pairs that, once mutated to cysteines, were likely to form disulfide bonds. This step led to the generation of 19 candidate enzymes, each with a different disulfide bond. We then simulated the WT and all 19 enzyme variants at three different temperatures (273, 310, and 340 K) and calculated the average RMSF at each simulated temperature. The slope ( $\lambda$ ) of the linear interpolation of avg-RMSF as a function of the temperature was used as a proxy for enzyme stability. For most mutants  $\lambda$  was larger than that of the WT, suggesting that the introduction of those disulfide bonds would not be effective in reducing the thermal vibration



**Figure 8.** Superposition of WT enzyme (gray) with the disulfide mutants (green). The SS03 mutant (Panel A) presents no significant alteration of any part of the enzyme due to the introduction of the disulfide bond, including the region near the double cysteine mutation (Panel C). The SS17 mutant (Panel B) presents no significant alteration in the overall structure of the enzyme and in the catalytic pocket. However, it shows a structural rearrangement close to the introduced mutations (Panel D) and a significant displacement of a loop (residues 63 to 69) that is far from the introduced mutation (orange circle in Panel B).

Mutant	RMSD [Å]
WT	0.99
SS03	0.93
SS17	1.35

**Table 5.** Root Mean Square Deviation (RMSD) comparison for native enzyme, and mutants SS03 and SS17. The RMSD is calculated, for each enzyme between the crystallographic structure and end-point of 5 ns long MD simulation at 300 K. The RMSD calculation was done considering main chain atoms only.

that could eventually lead to local and then global unfolding. Conversely, four mutant models (SS03, SS07, SS11, and SS17) featured a reduced  $\lambda$ , indicating a possible stabilizing effect associated with the designed mutations.

The four mutants selected based on our computational filter were then expressed in *E. coli*, providing yields similar to the WT except for SS11, which appeared to be very poorly expressed and showed no binding to the FAD cofactor (Fig. 4), suggesting a crucial role for residues K233 and/or P404 in the folding of the enzyme.

After incubation with fructosyl-lysine at increasing temperatures, SS07 appears to be slightly less active than the WT in the oxidized form, while showing stability levels comparable to those of the WT in the reduced form (Fig. 5, Panel C), suggesting a slightly detrimental effect associated with the introduction of the C106-C150 disulfide bond. Conversely, the SS03 mutant features a  $\sim 3^\circ\text{C}$  higher thermostability with respect to the WT (Fig. 5, Panel B). The SS17 variant, which shows an increase of  $\sim 8^\circ\text{C}$  in the  $T_{50}$  value and, contrary to all other tested mutants, a gradual decrease in residual activity at increasing temperatures, features detectable activity even after thermal treatment at  $95^\circ\text{C}$  (Fig. 5, Panel D). To confirm the formation of the designed disulfide bond in the SS03 and in the SS17 mutants and to assess any possible change that such modifications may have caused to their global molecular architecture relative to the wild-type enzyme, we solved the crystal structures of these two thermostabilized species. Both structures were determined at high resolution (2.15 and 2.85 Å, respectively). Globally, both SS03 and SS17 maintain the same fold as the wild-type (RMSD 0.268 and 0.804 Å, respectively). However, while the SS03 mutant superimposes almost perfectly with the wild type in all the regions, displaying only a very minor twist in the loop hosting one of the mutated residues (S67C), SS17 shows a remarkable conformational change involving the region where the aminoacids that were mutated to cysteines in the SS03 mutant (S67 and P121) are found, the distance between these two aminoacids in the SS17 structure being 11.4 Å vs. 5.0 Å in the structure of the wild type enzyme. This might explain why, contrary to our expectations, the combination of the two stabilizing disulfide bonds (SS03 and SS17) in the SSDM species does not provide any further stabilization



to the enzyme (Fig. 5, Panel E). However, this enzyme variant potentially carrying the double disulfide bonds appears to perform similarly to the SS03 variant, possibly suggesting that, in the course of the SSDM folding process, long range effects result in significantly pronounced deviations in the final enzyme conformation, thus likely allowing the formation of the C67-C121 disulfide bond while preventing the C295-C303 to form. Unfortunately, we could not obtain the structure of the SSDM species due to poor quality crystals. However, although we cannot be sure whether both disulfide bonds did indeed form in our double mutant variant, a feature that, as speculated, appears unlikely based on the crystal structure of the two single mutants that we have determined, our enzymatic data do rule out any possible cooperative effect associated with the simultaneous presence of these four specific cysteines in the sequence.

It should be noted that the proposed computational screening method produced two false positives: the SS11 variant does not fold under standard expression conditions, while SS07 shows a decreased enzymatic activity relative to the wild type form of the enzyme. In the case of SS11, we hypothesize that the mutations are in a critical position for the correct folding of the enzyme due to their proximity to the cofactor. Hence, they may affect the extremely conserved fold around the cofactor region. However, improper folding cannot be predicted with MD, and this is certainly a general limitation of any computational screening method. The SS07 variant, albeit correctly folded and active, shows reduced activity towards its natural substrate. In this case, the disulfide bond is located close to the entry tunnel and produces a narrowing effect that possibly hampers accessibility to the active site of the enzyme. In general, it is not surprising that our screening analysis produces a certain number of false positives. However, we believe that our computational screening method, especially in the initial phases of an enzyme design campaign, still provides a fast and efficient way to identify candidate active mutants while limiting the experimental burden and costs.

## Conclusions

Using a rational *in silico* computational design and screening method, we have produced two forms of the deglycating enzyme Amadoriase I featuring a significantly higher thermostability compared to the WT enzyme. In particular, the SS17 variant represents a possible starting point for the development of improved biosensors to be used for the detection of diabetes, for the design of therapeutic tools aimed at the prevention or the reduction of protein glycation in biological tissues, and as a potential tool for the reduction of acrylamide formation in food processing. The proposed computational screening method appears to be a promising tool to expedite the design of thermostable enzymes. Compared to directed evolution, the method used herein allows for a significant reduction in cost and time for the experimental production and characterization of candidate thermostable enzyme variants. Future work will be aimed at testing this methodology on other type of mutations beyond disulfide bonds, and extending the validation of our screening strategy to other enzymes.

Concerning glycated protein detection, stabilized enzymes may improve the typical shelf-life of the biosensors, of which the enzymes are usually the most sensitive component.

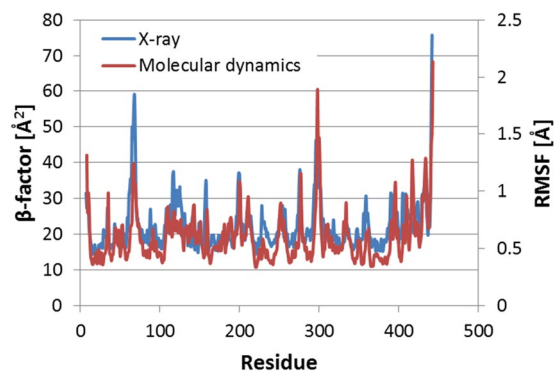
The application of Amadoriases in the food industry as a tool for acrylamide reduction requires that the enzyme be compatible with the high temperatures that are commonly associated with food treatments. High temperatures are known to facilitate the formation of Amadori products, which in some cases are undesired as they change the organoleptic as well as the nutritional profile of foods, often resulting in potentially harmful byproducts (as asparagine glycation, which turns into acrylamide). For example, roasting requires temperatures from 204 °C and above, frying in the range of 177–191 °C, while baking occurs at oven temperatures around 250 °C. On the other hand, there are some other thermal treatments, such as those related to milk conservation that take place at lower temperatures. For example, high-temperature short-time (HTST) milk pasteurization is done at 72 °C for 15 seconds, while milk ultra-heat treatment (UHT) is done at 140 °C (284 °F) for four seconds. Also, other treatments related to milk conservation occur within a temperature range from 63 °C for 30 minutes (the typical FDA-approved home pasteurization treatment) to 71.7–72 °C for 15 seconds (UK Dairy Products Hygiene Regulations 1995). For these applications, the increase in  $T_{50}$  shown by our mutants may well represent a step forward in the direction of obtaining a stable and useful enzyme for glycosylation prevention.

Finally, the use of Amadoriases to prevent protein glycation *in vivo* will require extensive mutagenesis and reduction of secondary structure elements to provide larger access to the catalytic site. Owing to the likely destabilization effect that is expected to be associated with such extensive structural modifications<sup>37</sup>, Amadoriases with enhanced stability will more likely represent more suitable candidates for the introduction of large yet sustainable structural changes in the enzyme.

## Material and Methods

**Computational design of disulfide bonds.** To identify suitable positions for the introduction of disulfide bonds, we used the disulfide bond prediction program SSBOND<sup>39</sup>. The 1.67 Å crystal structure of the wildtype (WT) enzyme Amadoriase I<sup>41</sup> (PDB code: 4XWZ) was used as a template. To avoid mutations that may affect enzymatic activity, we allowed only the substitution of those residues that are located farther than 10 Å from the FAD cofactor or from the ligand position. SSBOND allows the identification of potential disulfide bond sites based on conformational and energy constraints determined from disulfide bond-containing protein structures that are deposited in the PDB. Using an energy cut-off of 10 kcal/mol, the program generated a putative list of 19 pairs of residues that, if mutated to cysteines, would allow the introduction of stabilizing disulfide bonds in the structure of the enzyme. The output of the program includes the dihedral angles that describe the disulfide bond conformation and an energy rank that estimates the conformational strain thus introduced.

**Molecular Dynamics simulations.** We tested the 19 Amadoriase I variants (named SS01 to SS19), each with a different disulfide bond, using molecular dynamics (MD) simulations and we derived those putative sites where the introduction of an SS bond may enhance protein stability. The 19 molecular models were minimized



**Figure 9.** Superposition of the RMSF of WT-Amadoriase I and experimental  $\beta$ -factor, showing how the RMSF from MD simulations reproduces the same trend of the experimental  $\beta$ -factors and it allows to pinpoint the same peaks, which represent highly flexible protein regions.

and equilibrated following protocols used in previous studies<sup>44,58–60</sup>. Specifically, each variant was modelled using the AMBER99SBildn force field<sup>61</sup> for protein, water (TIP3P) and ions, while the FAD cofactor was modelled using the general amber force field (GAFF)<sup>62</sup>. The enzyme models were then solvated with  $\approx 14,992$  TIP3P water molecules. Since the enzyme carries a net charge ( $-10$  for the WT, while a slightly different one for each variant depending on the mutation), ions (specifically,  $\text{Na}^+$  and  $\text{Cl}^-$ ) were added to neutralize the system and provide a physiological ionic strength of  $0.15$  mol/L. The setup resulted in systems consisting of  $\approx 50,000$  atoms in a simulation box with initial dimensions of approximately  $70 \times 70 \times 85 \text{ \AA}^3$ .

Using NAMD code<sup>63</sup>, the systems were first minimized for 2,000 steps using the conjugate gradient algorithm. Then, they were equilibrated in NPT ensemble at 1 atm and at three different temperatures for each variant (273, 300, and 340 K, respectively). Langevin dynamics was used to maintain constant temperature with a coupling coefficient of  $1 \text{ ps}^{-1}$ , and pressure was maintained using the Nosé-Hoover method. In the NPT simulations, we used a time step of 2 fs, a non-bonded cut-off of  $9 \text{ \AA}$ , rigid bonds and particle-mesh Ewald long-range electrostatics. During minimization and NPT equilibration, the  $\text{C}_\alpha$  atoms of the protein were restrained by a  $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  spring constant to prevent protein conformation to be disrupted in the initial phase of system equilibration. Finally, the production runs were done using ACEMD<sup>64</sup> and a time step of 4 fs, while all the other parameters (non-bonded cut-off, and PME) were maintained as in the equilibration phase. We simulated each variant for 5 ns at the three different temperatures (273, 300, and 340 K) in explicit solvent. The Root Mean Square Displacement (RMSD) was monitored to confirm the stability of the simulations (see Fig. S4 in Supplementary Information).

**Computational screening of mutants.** In order to define a quantitative mutation screening criteria, we used the Root Mean Square Fluctuation (RMSF), which correlates with the  $\beta$ -factor from the X-ray crystal structures<sup>65</sup> (Fig. 9). The RMSF was calculated over the MD trajectory (excluding the first ns) for the  $\text{C}_\alpha$  of the protein from from residue 10 to 437 (we excluded the N- and the C-terminus of the protein due to their intrinsic high mobility).

The RMSF was then calculated at increasing temperatures (273, 300, and 340 K). As shown in Fig. 2, Panel A, the RMSF baseline increases with the temperature due to higher kinetic energy. However, the RMSF increases significantly more in specific regions of the protein, which may be assumed to represent the origin of local protein unfolding. For each temperature, we calculated the average RMSF (avg-RMSF) and we calculated the slope ( $\lambda$ ) of avg-RMSF vs. temperature (Fig. 2, Panel B and Fig. S3 of Supplementary Information). The value of  $\lambda$  represents an estimate of the increase of enzyme flexibility with the temperature. If an SS-variant showed a  $\lambda$  higher than the one calculated for the WT, the proposed SS-mutant was rejected. Otherwise, if  $\lambda$  was lower than that of WT, the enzyme variant was regarded as more thermostable than the WT and hence selected for experimental validation.

In order to test the effect of the simulation length, we performed 50 ns long MD simulations (at temperature of 273, 300 and 340 K) for the wild type enzyme and for the variants SS03, SS07, SS11, and SS17. We observed that all the structures reached stability within the first ns (see Fig. S5 in Supplementary Information), also at the highest simulated temperature. Furthermore, we compared the conformation at 300 K for the 5 ns MD simulations with the conformation obtained with 50 ns MD simulations, showing negligible differences (see Fig. S6 in Supplementary Information). The RMSD difference between the conformations after 5 ns and after 50 ns is approximately  $1 \text{ \AA}$  in all cases (see Table S1 of SI).

In addition, increasing the simulation length from 5 ns (Fig. S7A) to 50 ns (Fig. S7B) did not produce significant changes in the results in terms of  $\lambda$  parameter (see also Table S2). Moreover, while introducing extra computational costs, the inclusion of simulations at higher temperatures (e.g., 380 K) did not alter the results (see Fig. S7C).

**Materials.** DNA primers were from Metabion International AG. The gene sequence coding for Amadoriase I was cloned in a pET-3a vector (pET3a-AmadI) in our laboratory, and the cloning procedure has been described elsewhere<sup>41</sup>. Fructosyl-lysine synthesis and characterization is described in the Supplementary Information. All the chemicals were from Sigma-Aldrich, unless otherwise indicated, and were of the highest commercial purity available ( $>95\%$ ).

**Protein expression and purification.** The point mutations required for the selected double-cysteine mutants were introduced by PCR in the pET3a-Amad I using the QuikChange II site-directed mutagenesis kit (Agilent) according to the manufacturer's protocol. The generated mutants and the used primers are shown in Table S3 of Supplementary Information. The resulting mutant plasmids were analyzed and validated by DNA sequencing. All the proteins used in the present study were expressed in *E. coli* BL21(DE3)pLysS cells (Invitrogen) and purified following the same protocol. Cells were grown at 37 °C in 1 liter of Luria-Bertani (LB) broth, supplemented with 50 mg/liter ampicillin until they reached an  $OD_{600} = 0.6$ . After induction with isopropyl 1-thio- $\beta$ -D-galactopyranoside (IPTG) to a final concentration of 0.5 mM, the bacteria were cultured overnight at 25 °C. Cells were then harvested by centrifugation and resuspended in Buffer A (50 mM Tris-HCl pH 7.4, 150 mM NaCl) supplemented with a protease inhibitor cocktail and DNase and then lysed by sonication on ice. The lysate was clarified by centrifugation and the soluble fraction was loaded onto a Ni-NTA (Qiagen) column equilibrated with Buffer A. The Ni-NTA beads were first washed using 5 column volumes of Buffer A supplemented with 40 mM imidazole, then the bound N-terminal His-tagged enzyme was eluted with same buffer supplemented with 400 mM imidazole. The positive fractions were pooled, concentrated using an Amicon 20 centrifugal filter with a molecular weight cut-off of 10 kDa and loaded onto a Superdex 200 Increase 10/300GL size exclusion column (GE Healthcare) pre-equilibrated with Buffer B (10 mM Tris buffer pH 8.0). The resulting protein was collected, concentrated to  $\approx 10$  mg/mL and stored at  $-80$  °C. Protein concentration was assessed using a Bradford assay<sup>66</sup> kit (Bio-Rad) and bovine serum albumin (Sigma) as the standard. Sample purity was assessed by 10% SDS-PAGE.

**Absorption spectra.** All spectra were recorded at 25 °C in a solution of Buffer B, containing protein concentrations ranging between 1 and 10  $\mu$ M, with a Tecan Spark10 M and normalized to the absorbance at 280 nm.

**Enzyme activity assay.** Enzymatic activity was followed by a continuous assay that detects glucosone formation over time from fructosyl-lysine at 322 nm, as previously described<sup>41</sup>. The assay was adapted to support high throughput format in a 96 transparent polystyrene plates from Greiner Bio. The 200  $\mu$ L reaction mixture contained 10 mM Tris HCl pH 7.4, 20 mM o-phenylenediamine, 2 mM fructosyl-lysine. After 1 minute of pre-incubation, the reaction was started by adding the enzyme at a final concentration of 0.44  $\mu$ M, and the increase in absorbance at 322 nm (glucosone  $\epsilon_{322} = 149.25 \text{ M}^{-1} \text{ cm}^{-1}$ ) was monitored in a Spark10 M (Tecan). Unless otherwise indicated, enzymatic activity was assayed at 25 °C. One unit (U) is defined as the amount of enzyme required to produce 1  $\mu$ mol of glucosone per minute, and specific activity was expressed as  $\text{U mg}^{-1}$  of enzyme.

**Steady-State kinetics.** Apparent steady-state kinetics measurement for all the enzymes over its natural substrate were determined by means of the assay described above, with fructosyl-lysine concentrations varying from 0.05 to 2 mM. Data points were obtained from three independent experiments. Data were fitted by non-linear least-square fit of the data, with Eq. 1 (the Michaelis-Menten equation for hyperbolic substrate kinetics) using Hyperbola fit function of GraphPad Prism version 5.00 for Windows, GraphPad Software, La Jolla California USA.

$$v = \frac{V_{max} * S}{(K_m + S)} \quad (1)$$

in which  $v$ ,  $V_{max}$ ,  $S$ , and  $K_m$  represent the steady state reaction rate, maximum reaction rate, substrate concentration, and Michaelis-Menten constant for the substrate, respectively.

**Characterization of the thermoresistant mutants.** Thermal stability tests were performed using the assay described above after heat treatment, by incubating for 10 minutes the enzyme to the target temperature ranging from 25 to 100 °C (with 5 °C steps) in Buffer B, in the absence of ligands, and then cooling it down at 4 °C until test. The reduced forms of the enzymes were obtained by supplementing the buffer with 100 mM 1,4-dithiothreitol (DTT). After 1 h of incubation, the heat treatment and enzymatic assay were performed as for the oxidized forms. The  $T_{50}$  values were obtained by fitting the data with the Boltzmann equation (Eq. 2) using the Boltzmann sigmoidal fit function implemented in GraphPad Prism version 5.00 for Windows, GraphPad Software, La Jolla California USA.

$$A = A_{bottom} + \frac{(A_{top} - A_{bottom})}{1 + e^{\frac{T - T_{50}}{s}}} \quad (2)$$

where  $A$  represents the residual activity,  $A_{bottom}$  the lower asymptote of residual activity,  $A_{top}$  the higher asymptote of residual activity,  $T$  the temperature,  $T_{50}$  the temperature at which residual activity is halfway between  $A_{top}$  and  $A_{bottom}$ , and  $s$  the steepness of the curve.

The optimal pH for the WT and SS-mutants was determined by measuring the enzyme activity using the same assay from pH 6.0 to pH 11.0 with the following buffers (10 mM): Sodium Citrate (pH 6.0); Tris-HCl (pH 7.0, 8.0 and 9.0); CAPS-NaOH (pH 10.0, 11.0).

For all the listed experiments, data points were obtained from three independent experiments.

**Protein crystallization and Structure determination.** Crystals of both the SS03 and the SS17 mutant were obtained using the vapor diffusion method at room temperature by mixing a 1  $\mu$ L drop of  $\sim 15$  mg/mL protein

Crystal	SS03 (PDB id: 5OC3)	SS17 (PDB id: 5OC2)
<b>Data collection</b>		
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub>
Cell dimensions		
<i>a</i> (Å)	70.157	68.853
<i>b</i> (Å)	83.081	90.462
<i>c</i> (Å)	175.795	81.307
$\beta$ (°)		102.89
Wavelength (Å)	1	1
Resolution (Å)	47.87–2.15	48.82–2.85
$R_{\text{sym}}$ or $R_{\text{merge}}$ (%)	9.4 (31.1)	16.5 (49.5)
$I/\sigma I$	15.95 (6.3)	10.6 (4.1)
Completeness (%)	99.9 (99.7)	100.0(100.0)
Multiplicity	6.8	6.9
<b>Refinement</b>		
Resolution (Å)	47.87–2.15	46.37–2.85
No. of reflections	56474	22838
$R_{\text{work}}/R_{\text{free}}$ (%)	13.92/18.07	17.18/24.40
No. of atoms		
Protein	6935	6900
Ligand: FAD-FLY	106	106
Water	923	282
Average B-factors (Å <sup>2</sup> )		
Protein (chain A – chain B)	10.66–12.26	11.83–13.15
Ligand	5.17	7.31
Water	23.24	8.89
r-m-s-d		
Bond lengths (Å)	0.009	0.008
Bond angles (°)	1.173	1.23
Ramachandran		
Most favoured (%)	97.96	95.79
Additional allowed (%)	2.04	3.87
Disallowed (%)	0.0	0.34

**Table 6.** Data collection and refinement statistics of the SS03 and the SS17 crystal structures.

sample with an equal volume of a 0.1 M sodium citrate pH 5.6, 14% PEG4K, 15% isopropanol and 0.1 M sodium citrate pH 5.6, 14% Peg4K, 5% dimethyl sulfoxide solution respectively. Medium-size (150 × 100 × 50 μm) rod-like crystals appeared within a few days. Prior to X-ray data collection, crystals were frozen in a chemically identical solution supplemented with 25% (v/v) glycerol for cryo-protection. A 2.19 Å resolution data set and a 2.85 Å resolution data set were collected from a crystal of SS03 and a crystal of SS17, respectively, in both cases using  $\lambda = 1.000$  Å in the X06DA-PXIII beamline at the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland). Diffraction images were processed and scaled using XDS<sup>67</sup>. The structures were determined by molecular replacement using MOLREP<sup>68</sup> from the CCP4 package<sup>69</sup> and the free Amadoriase I structure (PDB code: 4WCT) as the search probe. Model building and refinement were carried out using REFMAC5<sup>70</sup> and PHENIX<sup>71</sup>. Water molecules were added both automatically using the phenix\_refine tool from the PHENIX package and manually from visual inspection of the electron density map. All the figures in the paper were generated using PyMOL<sup>45</sup> or VMD<sup>72</sup>. The refinement of the SS03 and the SS17 structures converged to a final  $R/R_{\text{free}} = 13.92/18.07\%$  and  $17.18/24.40\%$ , respectively. Data collection and final refinement statistics are shown in Table 6.

**Accession Numbers.** The final crystallographic coordinates of the crystal structures shown here are available in the RCSB PDB (accession code: 5OC2 and 5OC3).

## References

- Takahashi, M., Pischetsrieder, M. & Monnier, V. M. Isolation, purification, and characterization of amadoriase isoenzymes (fructosyl amine-oxidoreductase EC 1.5.3) from *Aspergillus* sp. *J. Biol. Chem.* **272**, 3437–3443 (1997).
- Ferri, S., Kim, S., Tsugawa, W. & Sode, K. Review of fructosyl amino acid oxidase engineering research: a glimpse into the future of hemoglobin A1c biosensing. *J. Diabetes Sci. Technol.* **3**, 585–592 (2009).
- Lin, Z. & Zheng, J. Occurrence, characteristics, and applications of fructosyl amine oxidases (amadoriases). *Appl. Microbiol. Biotechnol.* **86**, 1613–9 (2010).
- Wu, X. L., Palfey, B. A., Mossine, V. V. & Monnier, V. M. Kinetic studies, mechanism, and substrate specificity of amadoriase I from *Aspergillus* sp. *Biochemistry* **40**, 12886–12895 (2001).
- Gan, W. *et al.* Structural basis of the substrate specificity of the FPOD/FAOD family revealed by fructosyl peptide oxidase from *Eupenicillium terrenum*. *Acta Crystallogr. Sect. F, Struct. Biol. Commun.* **71**, 381–387 (2015).
- Miura, S., Ferri, S., Tsugawa, W., Kim, S. & Sode, K. Development of fructosyl amine oxidase specific to fructosyl valine by site-directed mutagenesis. *Protein Eng. Des. Sel.* **21**, 233–239 (2008).
- Kim, S., Miura, S., Ferri, S., Tsugawa, W. & Sode, K. Cumulative effect of amino acid substitution for the development of fructosyl valine-specific fructosyl amine oxidase. *Enzyme Microb. Technol.* **44**, 52–56 (2009).
- Miura, S., Ferri, S., Tsugawa, W., Kim, S. & Sode, K. Active site analysis of fructosyl amine oxidase using homology modeling and site-directed mutagenesis. *Biotechnol. Lett.* **28**, 1895–1900 (2006).
- Collard, F. *et al.* Crystal structure of the deglycating enzyme fructosamine oxidase (amadoriase II). *J. Biol. Chem.* **283**, 27007–16 (2008).
- Hatada, M. *et al.* Development of a screen-printed carbon electrode based disposable enzyme sensor strip for the measurement of glycated albumin. *Biosens. Bioelectron.* 1–7, <https://doi.org/10.1016/j.bios.2016.08.005> (2016).
- Monnier, V. M. & Wu, X. Enzymatic deglycation with amadoriase enzymes from *Aspergillus* sp. as a potential strategy against the complications of diabetes and aging. *Biochem Soc Trans* **31**, 1349–1353 (2003).
- Paul, R. G. & Bailey, A. J. Glycation of collagen: The basis of its central role in the late complications of ageing and diabetes. *Int. J. Biochem. Cell Biol.* **28**, 1297–1310 (1996).

13. Gautieri, A., Redaelli, A., Buehler, M. J. & Vesentini, S. Age- and diabetes-related nonenzymatic crosslinks in collagen fibrils: Candidate amino acids involved in Advanced Glycation End-products. *Matrix Biol.* **34**, 89–95 (2013).
14. Snedeker, J. G. & Gautieri, A. The role of collagen crosslinks in ageing and diabetes - the good, the bad, and the ugly. *Muscles. Ligaments Tendons J.* **4**, 303–308 (2014).
15. Sell, D. R. & Monnier, V. M. Molecular basis of arterial stiffening: role of glycation - a mini-review. *Gerontology* **58**, 227–37 (2012).
16. Del Turco, S. & Basta, G. An update on advanced glycation endproducts and atherosclerosis. *Biofactors* **38**, 266–274 (2012).
17. Goh, S. Y. & Cooper, M. E. The role of advanced glycation end products in progression and complications of diabetes. *J. Clin. Endocrinol. Metab.* **93**, 1143–1152 (2008).
18. Nagaraj, R. H., Linetsky, M. & Stitt, A. W. The pathogenic role of Maillard reaction in the aging eye. *Amino Acids* **42**, 1205–1220 (2012).
19. Vincent, A. M., Russell, J. W., Low, P. & Feldman, E. L. Oxidative stress in the pathogenesis of diabetic neuropathy. *Endocr. Rev.* **25**, 612–628 (2004).
20. Capuano, E. *et al.* Studies on the effect of Amadoriase from *Aspergillus fumigatus* on peptide and protein glycation in vitro. *J. Agric. Food Chem.* **55**, 4189–95 (2007).
21. Qian, Y., Zheng, J. & Lin, Z. L. Loop engineering of amadoriase II and mutational cooperativity. *Appl. Microbiol. Biotechnol.* **97**, 8599–8607 (2013).
22. Mottram, D. S., Wedzicha, B. L. & Dodson, A. T. Food chemistry: Acrylamide is formed in the Maillard reaction. *Nature* **419**, 448–449 (2002).
23. Stadler, R. H. *et al.* Acrylamide from Maillard reaction products. *Nature* **419**, 449–450 (2002).
24. Capuano, E. & Fogliano, V. Acrylamide and 5-hydroxymethylfurfural (HMF): A review on metabolism, toxicity, occurrence in food and mitigation strategies. *LWT - Food Sci. Technol.* **44**, 793–810 (2011).
25. Friedman, M. & Levin, C. E. Review of methods for the reduction of dietary content and toxicity of acrylamide. *J. Agric. Food Chem.* **56**, 6113–6140 (2008).
26. Gilis, D. & Rooman, M. PoPMuSiC, an algorithm for predicting protein mutant stability changes. *Application to prion proteins. Protein Eng. Des. Sel.* **13**, 849–856 (2000).
27. Yin, S., Ding, F. & Dokholyan, N. V. Eris: an automated estimator of protein stability. *Nat. Methods* **4**, 466–467 (2007).
28. Khan, S. & Vihinen, M. Performance of protein stability predictors. *Hum. Mutat.* **31**, 675–684 (2010).
29. Reetz, M. T., Carballeira, J. D. & Vogel, A. Iterative saturation mutagenesis on the basis of b factors as a strategy for increasing protein thermostability. *Angew. Chemie - Int. Ed.* **45**, 7745–7751 (2006).
30. Alfaro, P. *et al.* Optimization of designed armadillo repeat proteins by molecular dynamics simulations and NMR spectroscopy. *Protein Sci.* **21**, 1298–1314 (2012).
31. Lehmann, M., Pasamontes, L., Lassen, S. F. & Wyss, M. The consensus concept for thermostability engineering of proteins. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology* **1543**, 408–415 (2000).
32. Luke, K. A., Higgins, C. L. & Wittung-Stafshede, P. Thermodynamic stability and folding of proteins from hyperthermophilic organisms. *FEBS Journal* **274**, 4023–4033 (2007).
33. Borgo, B. & Havranek, J. J. Automated selection of stabilizing mutations in designed and natural proteins. *Proc. Natl. Acad. Sci.* **109**, 1494–1499 (2012).
34. Jacak, R., Leaver-Fay, A. & Kuhlman, B. Computational protein design with explicit consideration of surface hydrophobic patches. *Proteins Struct. Funct. Bioinforma.* **80**, 825–838 (2012).
35. Magliery, T. J. Protein stability: Computation, sequence statistics, and new experimental methods. *Current Opinion in Structural Biology* **33**, 161–168 (2015).
36. Khersonsky, O. *et al.* Optimization of the in-silico-designed Kemp eliminase KE70 by computational design and directed evolution. *J. Mol. Biol.* **407**, 391–412 (2011).
37. Sullivan, B. J. *et al.* Stabilizing proteins from sequence statistics: The interplay of conservation and correlation in triosephosphate isomerase stability. *J. Mol. Biol.* **420**, 384–399 (2012).
38. Trudeau, D. L., Lee, T. M. & Arnold, F. H. Engineered thermostable fungal cellulases exhibit efficient synergistic cellulose hydrolysis at elevated temperatures. *Biotechnol. Bioeng.* **111**, 2390–2397 (2014).
39. Hazes, B. & Dijkstra, B. W. Model building of disulfide bonds in proteins with known three-dimensional structure. *Protein Eng. Des. Sel.* **2**, 119–125 (1988).
40. Hazes, B. & Dijkstra, B. W. SSBOND. <http://hazeslab.med.ualberta.ca/programs/ssbond.ht> (1998).
41. Rigoldi, F. *et al.* Crystal structure of the deglycating enzyme amadoriase i in its free form and substrate-bound complex. *Proteins* **84**, 744–758 (2016).
42. Mennella, C., Borrelli, R. C., Vinale, F., Ruocco, M. & Fogliano, V. Substrate specificity of amadoriase I from *Aspergillus fumigatus*. *Ann. N. Y. Acad. Sci.* **1043**, 837–44 (2005).
43. Chovancova, E. *et al.* CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *Plos Comput. Biol.* **8**, (2012).
44. Rigoldi, F. *et al.* Molecular dynamics simulations provide insights into the substrate specificity of FAOX family members. *Mol. Biosyst.* **12**, 2622–2633 (2016).
45. Schrödinger. Pymol. <https://pymol.org/2> (2017).
46. Guo, C. *et al.* A novel strategy for thermostability improvement of trypsin based on N-glycosylation within the ??-loop region. *J. Microbiol. Biotechnol.* **26**, 1163–1172 (2016).
47. Yao, M. Z., Wang, X., Wang, W., Fu, Y. J. & Liang, A. H. Improving the thermostability of *Escherichia coli* phytase, appA, by enhancement of glycosylation. *Biotechnol. Lett.* **35**, 1669–1676 (2013).
48. Sanchez-Romero, I. *et al.* Mechanism of Protein Kinetic Stabilization by Engineered Disulfide Crosslinks. *PLoS One* **8**, e70013 (2013).
49. Jermutus, L., Tessier, M., Pasamontes, L., Van Loon, A. P. G. M. & Lehmann, M. Structure-based chimeric enzymes as an alternative to directed enzyme evolution: Phytase as a test case. *J. Biotechnol.* **85**, 15–24 (2001).
50. Zhang, W., Mullaney, E. J. & Xin, G. L. Adopting selected hydrogen bonding and ionic interactions from *Aspergillus fumigatus* phytase structure improves the thermostability of *Aspergillus niger* PhyA phytase. *Appl. Environ. Microbiol.* **73**, 3069–3076 (2007).
51. Betz, S. F. Disulfide bonds and the stability of globular proteins. *Protein Sci.* **2**, 1551–8 (1993).
52. Abkevich, V. I. & Shakhnovich, E. I. What can disulfide bonds tell us about protein energetics, function and folding: simulations and bioinformatics analysis. *J. Mol. Biol.* **300**, 975–985 (2000).
53. Goldenzweig, A. *et al.* Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Mol. Cell* **63**, 337–346 (2016).
54. Turner, N. J. Directed evolution drives the next generation of biocatalysts. *Nat. Chem. Biol.* **5**, 567–573 (2009).
55. Bommaris, A. S., Broering, J. M., Chaparro-Riggers, J. F. & Polizzi, K. M. High-throughput screening for enhanced protein stability. *Current Opinion in Biotechnology* **17**, 606–610 (2006).
56. Sakaue, R. & Kajiyama, N. Thermostabilization of bacterial fructosyl-amino acid oxidase by directed evolution. *Appl. Environ. Microbiol.* **69**, 139–145 (2003).
57. Tokuriki, N., Stricher, F., Serrano, L. & Tawfik, D. S. How protein stability and new functions trade off. *PLoS Comput. Biol.* **4**, (2008).

58. Nair, A. K., Gautieri, A. & Buehler, M. J. Role of Intrafibrillar Collagen Mineralization in Defining the Compressive Properties of Nascent Bone. *Biomacromolecules* **15**, 2494–500 (2014).
59. Gautieri, A., Vesentini, S. & Redaelli, A. How to predict diffusion of medium-sized molecules in polymer matrices. from atomistic to coarse grain simulations. *J. Mol. Model.* **16**, 1845–1851 (2010).
60. Gautieri, A. *et al.* Computer-Aided Molecular Modeling and Experimental Validation of Water Permeability Properties in Biosynthetic Materials. *J. Comput. Theor. Nanosci.* **7**, 1287–1293 (2010).
61. Case, D. A. *et al.* The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **26**, 1668–1688 (2005).
62. Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
63. Theoretical and Computational Biophysics Group, U. NAMD 2.13. <http://www.ks.uiuc.edu/Research/namd> (2017).
64. Acellera. ACEMD. <https://www.acellera.com/products/molecular-dynami> (2017).
65. Wijma, H. J. *et al.* Computationally designed libraries for rapid enzyme stabilization. *Protein Eng. Des. Sel.* **27**, 49–58 (2014).
66. Hammond, J. B. & Kruger, N. J. The Bradford method for protein quantitation. *Methods Mol. Biol.* **3**, 25–32 (1988).
67. MPI for Medical Research. XSD, X-ray Detector Software. <http://xds.mpimf-heidelberg.mpg.de> (2007).
68. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. Sect. D-Biological Crystallogr.* **66**, 22–25 (2010).
69. CCLRC Daresbury Laboratory. CCP4. <http://www.ccp4.ac.uk> (2017).
70. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D-Biological Crystallogr.* **67**, 355–367 (2011).
71. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D-Biological Crystallogr.* **66**, 213–221 (2010).
72. Theoretical and Computational Biophysics Group, U. VMD 1.9.3. <http://www.ks.uiuc.edu/Research/vmd> (2017).

## Acknowledgements

The authors wish to thank Istituto Italiano di Tecnologia, Fondazione Cariplo (grant 2016–0481) and H2020 EU Project AMMODIT (Project ID 645672) for funding. They also thank the X06DA-PXIII beamline personnel of the Paul Scherrer Institute, Villigen, Switzerland, for help with the data collection. The authors declare that they have no conflicts of interest with the contents of this article.

## Author Contributions

A.G., E.P., F.R., S.D. designed the research. F.R., A.G. performed the computational simulations, F.R., A.G. and A.R. analyzed the data. S.D. F.R., A.G., F.G., and F.S. performed the experiments, S.D. F.R., A.G., F.G., F.S., E.P. and T.B. analyzed the data. All authors wrote and reviewed the paper.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-19991-x>.

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

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018