Received: 18 May 2020 | Revised: 27 July 2020 | Accepted: 9 August 2020

DOI: 10.1002/bit.27535

ARTICLE

BIOTECHNOLOGY AND BIOENGINEERING WILEY

Rational backbone redesign of a fructosyl peptide oxidase to widen its active site access tunnel

Federica Rigoldi[1] | Stefano Donini[2] | Archimede Torretta[2] | Anna Carbone[3,4] | Alberto Redaelli[1] | Tiziano Bandiera[4] | Emilio Parisini[2,5] | Alfonso Gautieri[1]

[1]Dipartimento di Elettronica, Informazione e Bioingegneria, Biomolecular Engineering Lab, Politecnico di Milano, Milano, Italy

[2]Center for Nano Science and Technology@Polimi, Istituto Italiano di Tecnologia, Milano, Italy

[3]Department of Biological, Chemical and Pharmaceutical Sciences and Technologies (STEBICEF), University of Palermo, Palermo, Italy

[4]D3-PharmaChemistry, Istituto Italiano di Tecnologia, Genova, Italy

[5]Biotechnology Group, Latvian Institute of Organic Synthesis, Riga, Latvia

Correspondence
Emilio Parisini, Center for Nano Science and Technology@Polimi, Istituto Italiano di Tecnologia, Via G. Pascoli 70/3, 20133 Milano, Italy.
Email: emilio.parisini@iit.it and emilio.parisini@osi.lv

Alfonso Gautieri, Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy.
Email: alfonso.gautieri@polimi.it

Funding information
Fondazione Cariplo, Grant/Award Number: 2016-0481

Abstract

Fructosyl peptide oxidases (FPOXs) are enzymes currently used in enzymatic assays to measure the concentration of glycated hemoglobin and albumin in blood samples, which serve as biomarkers of diabetes. However, since FPOX are unable to work directly on glycated proteins, current enzymatic assays are based on a preliminary proteolytic digestion of the target proteins. Herein, to improve the speed and costs of the enzymatic assays for diabetes testing, we applied a rational design approach to engineer a novel enzyme with a wider access tunnel to the catalytic site, using a combination of Rosetta design and molecular dynamics simulations. Our final design, L3_35A, shows a significantly wider and shorter access tunnel, resulting from the deletion of five-amino acids lining the gate structures and from a total of 35 point mutations relative to the wild-type (WT) enzyme. Indeed, upon experimental testing, our engineered enzyme shows good structural stability and maintains significant activity relative to the WT.

KEYWORDS

access tunnel, biosensor, diabetes, fructosyl peptide oxidase, rational enzyme design

1 | INTRODUCTION

Fructosyl amino acid oxidases (FAOX; Ferri, Kim, Tsugawa, & Sode, 2009; Lin & Zheng, 2010; Takahashi, Pischetsrieder, & Monnier, 1997) are a family of fungal and bacterial enzymes that cleave low-molecular-weight Amadori product (i.e., fructosyl amino acids) to generate a free amine, glucosone, and hydrogen peroxide (Gan et al., 2015; Rigoldi, Spero et al., 2016; Wu, Palfey, Mossine, & Monnier, 2001). Among these enzymes, those that also exhibit reactivity for small fructosyl peptides (up to six amino acids; Ferri, Miyamoto, Sakaguchi-Mikami, Tsugawa, & Sode, 2013) are specifically termed fructosyl peptide oxidases (FPOX). Today, FPOX are used for the detection of glycated hemoglobin (HbA1c), a long-term marker for diabetes (the half-life of HbA1c is 3 months; Kim, Miura, Ferri, Tsugawa, & Sode, 2009; Miura, Ferri, Tsugawa, Kim, & Sode, 2006; Miura, Ferri, Tsugawa, Kim, & Sode, 2008). Current HbA1c enzymatic assays are based on a preliminary proteolytic digestion of HbA1c, which releases the glycated N-terminal dipeptide, fructosyl valine histidine (fVH). Then, the FPOX enzyme binds and hydrolyzes the released fVH dipeptide, thus producing hydrogen

Federica Rigoldi and Stefano Donini contributed equally to this study.

peroxide which, in turn, is measured in a colorimetric assay using horseradish peroxidase and a suitable chromophore (Collard et al., 2008). Via a similar mechanism, FPOX enzymes can be used for the detection of glycated albumin, a short-to mid-term glycemic marker for diabetes (glycated albumin has a half-life of 3 weeks; Shahbazmohammadi, Sardari, Lari, & Omidinia, 2019). To address the rapidly growing demand for more convenient monitoring of diabetes mellitus, efforts are aimed at developing enzymatic assays for glycated hemoglobin and albumin that do not require the preliminary proteolytic digestion of the target proteins. Indeed, the elimination of the proteolytic stage would substantially contribute to the simplification of the assay, to the reduction of its costs, and to the shortening of the measuring time. To the best of our knowledge, the only currently known enzyme that is capable of detecting HbA1c directly, that is, without a proteolysis step, has been recently obtained by introducing a single point mutation (R61G), to open the "gate" leading to the active pocket, followed by several random mutagenesis cycles to obtain a mutant showing catalytic activity on HbA1c (AnFPOX-21; Ogawa et al., 2019).

Besides their role in diabetes monitoring, FPOX family members are also regarded as promising therapeutic tools for the prevention or the reduction of protein glycation in biological tissues (Monnier & Wu, 2003). Glycation is the spontaneous, nonenzymatic, and irreversible reaction between a sugar moiety and a protein, leading to a covalent adduct, the Amadori product (Paul & Bailey, 1996). By modifying the chemistry of functional proteins (Gautieri, Redaelli, Buehler, & Vesentini, 2013; Snedeker & Gautieri, 2014), glycation leads to a cascade of adverse clinical outcomes including arterial stiffening (Sell & Monnier, 2012), atherosclerosis (Del Turco & Basta, 2012), nephropathy (Goh & Cooper, 2008), retinopathy (Nagaraj, Linetsky, & Stitt, 2012), and neuropathy (Vincent, Russell, Low, & Feldman, 2004). So far, the use of FPOX to prevent protein glycation is precluded because these enzymes show no significant activity on intact proteins (Capuano et al., 2007; Qian, Zheng, & Lin, 2013), which is due to their buried active site and to the narrow tunnel that provides access to their catalytic pocket, as shown by the crystal structures of FPOX enzymes (Rigoldi, Gautieri et al., 2016).

All the aforementioned applications of FPOX could largely benefit from the development of an improved enzyme that is capable of reacting with full-length glycated proteins. In this study, we apply a rational design approach to engineer a novel enzyme with a wider access tunnel to the catalytic site. Enzyme access tunnel engineering is a fast-growing field based on the evidence that this kind of modification can improve activity, specificity, promiscuity, enantioselectivity, and stability (Kokkonen, Bednar, Pinto, Prokop, & Damborsky, 2019). In this study, we compared the crystal structure of Amadoriase I (PDB ID 4WCT; Rigoldi, Gautieri et al., 2016; Rigoldi et al., 2018), which has been previously solved in our lab, with the structures of PnFPOX (PDB ID 5T1E; Shimasaki, Yoshida, Kamitori, & Sode, 2017) and of EtFPOX (PDB ID 4RSL; Gan et al., 2015) to identify conserved structural elements (Figure S1). Starting from the minimal conserved structure, we further removed one turn in the helix defining the access to the catalytic site. This preliminary template was then optimized using several cycles of Rosetta-based design (Kaufmann, Lemmon, Deluca, Sheehan, & Meiler, 2010) and molecular dynamics (MD) simulations to obtain a new, stable, and active FPOX enzyme that features a wider tunnel to the catalytic pocket.

2 | MATERIAL AND METHODS

2.1 | Structural superimposition and target identification for design

To identify the most suitable regions for structural modifications, we aligned the sequence and superimposed the structure of three different FPOX: Amadoriase I (PDB ID 4WCT; Rigoldi, Gautieri et al., 2016), PnFPOX (PDB ID 5T1E; Shimasaki et al., 2017), and EtFPOX (PDB ID 4RSL; Gan et al., 2015). The sequence alignment (Figure S1) was done using Clustal Omega software (Sievers et al., 2011), while the structural alignment was obtained with Pymol software (Schrödinger, 2017). For comparison, residue numbering is referred to as the 5T1E structure. Since 5T1E features both a shorter helix and a shorter loop relative to 4WCT and since its structure was solved at a higher resolution than 4RSL, we selected 5T1E as the basic scaffold for further modification steps.

2.2 | Scaffold remodeling

Starting from the crystal structure of 5T1E, we selected the eight-residue-long region (green box in Figure S1) corresponding to residues ranging from 62 to 69, and we replaced it with a shorter connecting sequence that we designed using a missing loop (Mandell, Coutsias, & Kortemme, 2009) coupled with remodel (Huang et al., 2011) functions.

The loop modeling function was applied to close the gap of the removed backbone segment and to rebuild it using two, three, or four residues, respectively. Since during loop reconstruction, the start and the endpoints are fixed, we applied the kinematic closure protocol. Besides loop modeling, we remodeled the 6 Å surrounding region (i.e., the portion that accounts for those interactions that were lost upon truncation), while freezing all those residues that are within 4 Å from the flavin adenine dinucleotide (FAD) cofactor. Repack for an outer shell of 8 Å around the designable region was also allowed. The design runs were done using the Rosetta remodel function. We used a blueprint file to specify those positions that could be designed, those that are subject to repack and those that are frozen, respectively. Finally, just before the beginning of the following α-helix, we replaced Q70 with a proline, a conserved amino acid in all three wild-type (WT) structures.

For each loop design, we generated 50 different models, producing a final pool of 150 possible candidates. All the output models were ranked based on the final total score (score 15) and the lowest five models for each loop length (i.e., a total of 15 models) were selected for further optimization steps.

2.3 | Loop length selection

Each of the 15 candidate models thus selected was subjected to a 50 ns MD simulation following protocols used in previous studies (Gautieri et al., 2010; Gautieri, Vesentini, Redaelli, & Buehler, 2009; Rigoldi et al., 2018). Specifically, we used the AMBER14SB force field (Case et al., 2005) for protein, water (TIP3P), and ions, while the general amber force field (Wang, Wolf, Caldwell, Kollman, & Case, 2004) was used for modeling the FAD cofactor. The 15 molecular models, solvated with a 15 Å pad of TIP3P water (resulting in a final simulation box of ≈50.000 atoms), were subjected to 1,000 energy minimization steps and were equilibrated at a pressure of 1 atm and at a temperature of 300 K with NAMD software (Nelson et al., 1996). In the NPT simulations, we used a time step of 2 fs, a nonbonded cutoff of 12 Å, rigid bonds, and particle-mesh Ewald (PME) long-range electrostatics. During these steps, the Cα atoms of the protein were restrained by a 10 kcal·mol^{-1}·Å^{-2} spring constant. Finally, the production runs were done using ACEMD (Harvey, Giupponi, & De Fabritiis, 2009), whereby all the parameters (nonbonded cutoff and PME) were fixed as in the equilibration phase, except for the time step, which was increased to 4 fs, with a final simulation time of 50 ns. Root mean square deviations (RMSD) were monitored during the MD runs, and for each system, the root mean square fluctuations (RMSFs) and structural alignment with the WT form of 5T1E were calculated to select the best candidate for the following steps.

2.4 | PROSS optimization

Structural and stability analysis suggested that the systems featuring a three-residue loop represent the most suitable candidates for experimental production. We also made use of the PROSS server (Goldenzweig et al., 2016) to identify possible point mutations that may increase the stability of the newly engineered protein. As input, we used the best candidate isolated from the previous structural analysis, and we excluded the residues that are within 5 Å from the FAD cofactor from being modified.

PROSS output consisted of a list of seven possible models featuring an increasing number of point mutations. The seven candidate variants were investigated with MD simulations, using the same protocol described in the previous section. Finally, we selected the model that featured the smallest RMSD relative to the starting 5T1E structure as the best candidate for experimental production.

2.5 | Protein expression and purification

The L3_35A mutant was expressed in *Escherichia coli* BL21 Star (DE3) cells (Invitrogen). Cells were grown at 37°C in 1L of Luria–Bertani broth, supplemented with 50 mg/L ampicillin until they reached an OD_{600} = 0.6. After induction with isopropyl 1-thio-β-D-galactopyranoside at a final concentration of 0.4 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 8.0, 150 mM NaCl, 0.5 mM FAD, and 5% glycerol) 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 five column volumes of buffer A supplemented with 40 mM imidazole, then the bound N-terminal His-tagged enzyme was eluted with the same buffer supplemented with 400 mM imidazole. The positive fractions were pooled, concentrated using an Amicon 20 centrifugal filter with a molecular cutoff of 10 KDa, and loaded onto a Superdex 200 Increase 10/300GL size exclusion column (GE Healthcare) pre-equilibrated with buffer A. The resulting protein was collected, concentrated to ≈5 mg/ml and stored at −80°C. Protein concentration was assessed using the Bradford Assay (Hammond & Kruger, 1988) Kit (Bio-Rad) and bovine serum albumin (Sigma-Aldrich) as the standard. Sample purity was assessed by 10% sodium dodecyl sulphate-polyacrylamide gel electrophoresis.

2.6 | Enzyme activity assay

Enzymatic activity was assayed following a previously reported protocol (Rigoldi et al., 2018). The assay indirectly quantifies the amount of glucosone formation over time from fructosyl-lysine at 322 nm. The 200 μl reaction mixture contained 10 mM Tris-HCl pH 7.4, 20 mM o-phenylenediamine, and 2 mM fructosyl-lysine. After 1 min of preincubation, the reaction was started by adding the enzyme at a final concentration of 0.44 μM, and the increase in absorbance at 322 nm (glucosone ε322 = 149.25 M^{-1}·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 μmol of glucosone per minute, and specific activity was expressed as U/mg of enzyme.

2.7 | Steady-state kinetics

Apparent steady-state kinetics measurement over fructosyl-lysine was determined by means of the same assay described above, with substrate concentrations varying from 0.05 to 5 mM. Data points are the result of three independent experiments. Data were fitted by nonlinear least square fit of the data, with Equation (1) (the Michaelis–Menten equation for hyperbolic substrate kinetics) using Sigma Plot (Systat Software, San Jose, CA)

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

where 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.

2.8 | Protein crystallization and structure determination

Initial crystallization screening was performed using commercial screens from Hampton Research with the sitting-drop vapor-diffusion method. In general, 2 µl of purified protein solution (≈14 mg/ml) was mixed with an equal volume of reservoir solution in 24-well Cryschem plates and was equilibrated against 500 µl of the reservoir solution at room temperature. Crystals of the L3_35A mutant were obtained within 1 day using Crystal screen 2 (Hampton Research) Condition No. 2: 0.1 M MES monohydrate pH 6.5, 12% polyethylene glycol 20 000. Before X-ray data collection, crystals were frozen in a chemically identical solution supplemented with 25% (vol/vol) glycerol for cryoprotection. A 1.38 Å resolution dataset was collected using $\lambda = 1.000$ Å in the X06DA-PXIII Beamline at the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland). Diffraction images were processed with iMosflm (Battye, Kontogiannis, Johnson, Powell, & Leslie, 2011), scaled using SCALA from the CCP4 program suite (Collaborative Computational Project. Number 4, 1994) and the crystal was assigned to space group $P4_22_12$ with unit cell parameters $a = 90.3$ Å, $b = 90.3$ Å, $c = 131.95$ Å, $\alpha = \beta = \gamma = 90°$. The asymmetric unit contains one protein molecule. The structure was determined by molecular replacement using program PHASER (McCoy et al., 2007) from the PHENIX program suite (Adams et al., 2010) and the structure of the FPOX from *Phaeospaeria nodrum* as the search model (PDB 5T1E). Model building was performed using AUTOBUILD (Terwilliger et al., 2007) and refinement were carried out using REFMAC5 (Murshudov et al., 2011), and PHENIX (Adams et al., 2010). 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. Coordinates and structure factors for the L3_35A mutant have been deposited in the Protein Data Bank with the PDB code 6Y4J. All the figures in the paper were generated using PyMOL (DeLano, 2002). The refinement of the L3_35A structure converged to a final $R/R_{free} = 14.88/16.91\%$. The tunnel dimension was analyzed with CAVER 3.0.3 software (Chovancova et al., 2012) using a probe radius of 1.4 Å, a shell depth of 4 Å, and a shell radius of 3 Å.

3 | RESULTS AND DISCUSSION

3.1 | Identification of design target

The measure of the backbone RMSD among the three structures (Table S1) suggests a shared global fold. Structure superimposition (colored by residues conservation; Figure 1a) confirms a common tridimensional arrangement with a distribution of the conserved residues over the entire sequences, especially in the key secondary structure motifs that define the Flavin cofactor pocket. The three enzymes present a common highly conserved central core (Figure 1a, red and orange regions), surrounded by variable regions that most likely confer substrate specificity (Figure 1a, blue regions).

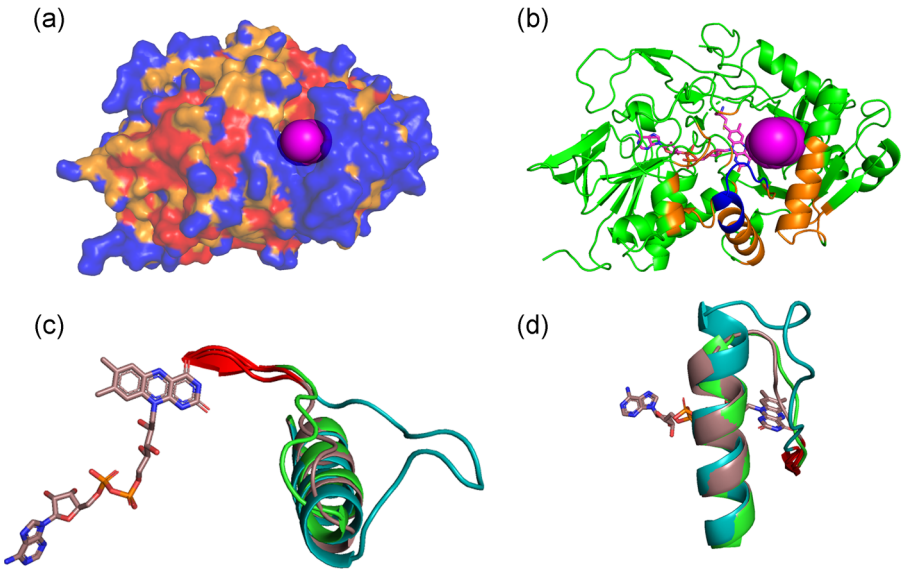

FIGURE 1 Structure of wild-type fructosyl peptide oxidases (FPOXs). (a) The structural similarity between 4WCT, 4RSL, and 5T1E. Surface representation of nonconserved (blue), strictly conserved (red), and homologous (orange) amino acids is shown. The three structures, belonging to the same enzyme family, share a core-common folding surrounded by variable-flanking regions. The tunnel leading to the active site is represented by purple spheres (identified using Caver 3.0.1). (b) The tunnel (represented as purple spheres) leading to the catalytic pocket and the surrounding regions. The backbone segment that lines the access tunnel and that was redesigned in this study is shown as a blue cartoon. The orange region shows those residues that are within 5 Å from the engineered segment, whose side chain were allowed to mutate. (c) The longer loop region found in the 4WCT structure compared to the other two structures. (d) The additional helix turn for 4WCT that is missing in the other two FPOX enzymes [Color figure can be viewed at wileyonlinelibrary.com]

Concerning the structures delimiting the tunnel to the active pocket, structural superimposition shows that these are the most variable regions of the enzymes. The largest structure delimiting the tunnel is a 17–22 amino acid alpha helix (α3 in Figure S1). This structure, corresponding to residue G100 to D120 in the 5T1E structure, is conserved in all three enzymes, although with some differences in the sequences, as shown in the alignment (Figure S1). Conversely, a second helical structure (residues P66 to L69) that, together with a loop region (residues G60 to N65), delimits the tunnel is significantly different in the three enzymes, both in sequence and in structure (Figure S1, green box). On the basis of these observations, we decided to redesign the 5T1E structure.

Structural analysis reveals that the region corresponding to residues K57-I58-M59, which lays within 5 Å from the FAD cofactor, is conserved in all three enzymes (red in Figure 1c). From residue 60 onwards, the three structures become highly divergent. In 4WCT, a loop of eight residues (Figure 1c) is followed by an eight-residue-long helix (corresponding to two α helix turns), while 5T1E and 4RSL feature a shorter loop and a shorter helix (Figure 1d).

The analysis of the WT forms shows a common L-shaped tunnel with an averaged bottleneck of 1.455 Å that leads to the active pocket (Figure 2a). The designable region (residues 62–69) acts as the access gate, restricting the tunnel and defining one of its sides. To enhance accessibility to the catalytic pocket, we removed this portion and we remodeled it with loops made of two, three, or four residues, respectively (Figure 2b).

3.2 | Loop design and selection

After the loop closure and remodeling phase, for each of the three loop-length families, we chose the five designs that showed the best Rosetta energy total score for further screening with MD simulations.

The analysis of the RMSD calculated over the whole protein backbone shows that all the 15 models reached stability in the first few nano seconds of simulation (Figure 3a–c), with no apparent difference depending on the loop length. The only exception is model L2_34, where the RMSD increases towards the end of the MD simulation.

RMSF analysis (Figure 3d–f) of the designed region and of the flanking regions (residues 55–75) shows improved local stability (with respect to the WT) for all models with loops made of three amino acids. Conversely, those designs in which shorter or longer loops were used, did not generally yield any improved local stability.

On the basis of RMSF analysis, we focused on the three-residue-long loop family for further rounds of design revisions. For each of the five different variants obtained at this stage, we compared the conformation at the end of the MD simulation with the WT structure. We chose the model that showed the lowest RMSD relative to the WT structure (i.e., model L3_35) as the candidate for further design.

Globally, at this stage of design, we substituted residues 62–69 with a three amino-acid-long loop, in addition to a single point mutation of position 70 into proline, for properly capping the shorter helix.

3.3 | Protein stability optimization

Model L3_35 was optimized using the PROSS online web server (Goldenzweig et al., 2016), keeping all those residues that are within 4 Å from the FAD cofactor frozen. Furthermore, we chose to fix also the catalytic residues to prevent their mutations in the course of the design process. PROSS produced seven enzyme variants (named from L3_35A to L3_35G), each featuring a different number of single point mutations that were aimed at improving protein solubility and stability. The PROSS designs were subjected to 50 ns of MD simulation, which showed that the global stability of the seven PROSS-generated candidate mutants was comparable to the WT enzyme, except for

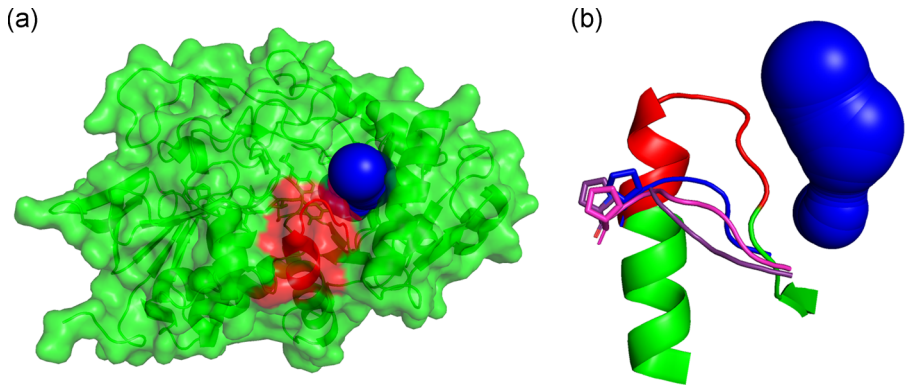

FIGURE 2 Design strategy. (a) The highlighted red region delimits the tunnel leading to the catalytic site and prevents the access of larger substrates. (b) The section of the wild-type enzyme that was removed (red), in comparison with the three families of connecting loop adopted in our approach: loops made of two amino acids (blue), loops made of three amino acids (pink), and loops made of four amino acids (purple). A proline, which is intended to act as a helix-capping residue, was always introduced in all the design attempts. The tunnel is shown as blue spheres [Color figure can be viewed at wileyonlinelibrary.com]

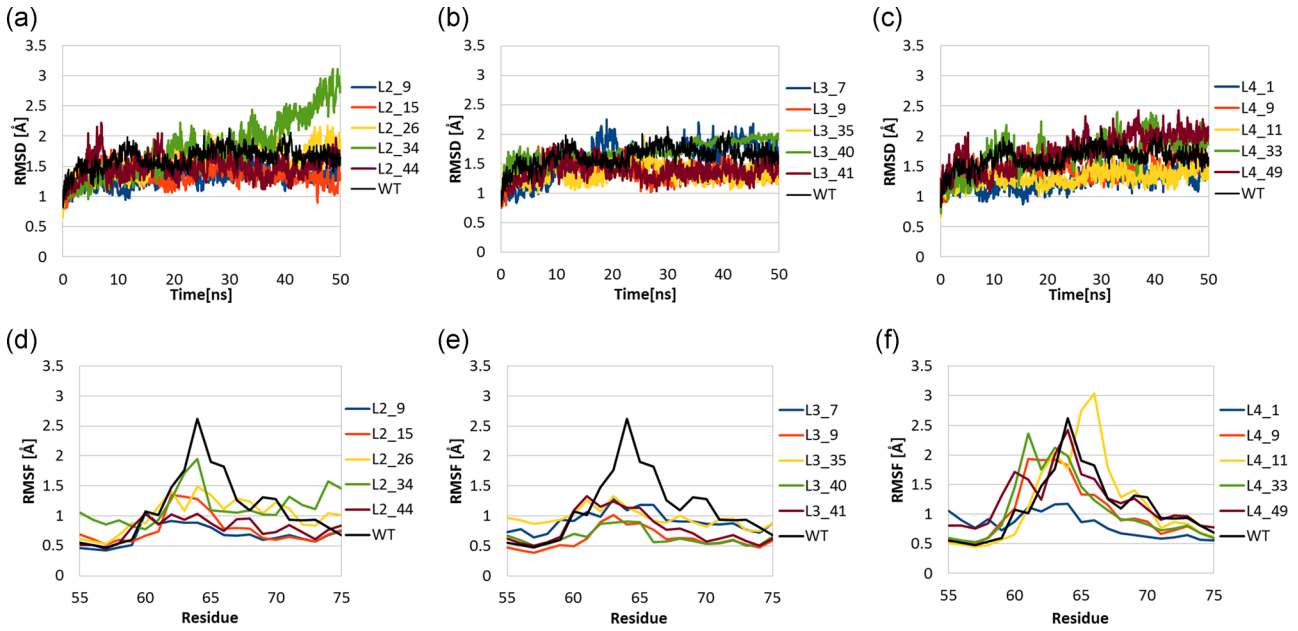

FIGURE 3 RMSD analysis. Global stability analysis shows no significant difference between the wild-type and the models that were designed with loops of two residues (a), three residues (b), or four residues (c). All the models reach global stability within a few nanoseconds. RMSF analysis of the designed region and flanking regions (residues from 55 to 75) shows improved stability when the designed loop is made of three residues (e), while mixed results are obtained with loops of two (d) or four amino acids (f). RMSD, root mean square deviation; RMSF, root mean square fluctuation [Color figure can be viewed at wileyonlinelibrary.com]

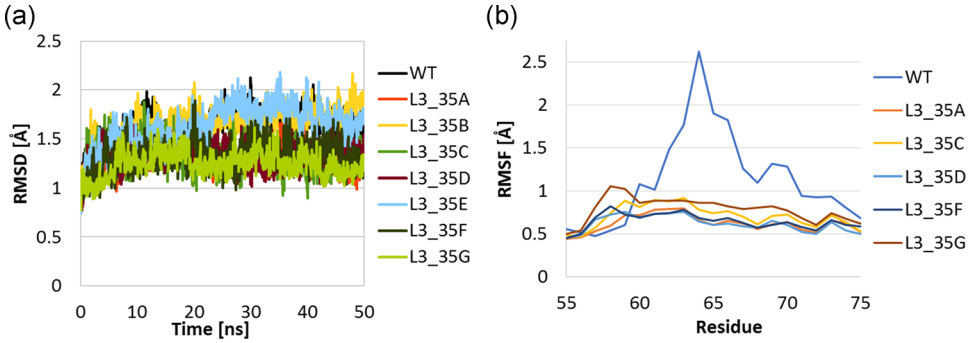

FIGURE 4 RMSD and RMSF comparison for the seven designed enzyme candidates proposed after PROSS optimization. For these positions, we observe an RMSD of 0.70 Å for backbone atoms (Figure 7b), suggesting that no change occurs in the local fold was averaged on all the backbone atoms of the structures, while RMSF evaluation was restricted to the designed region and to the flanking amino acids (residues 55–75). (a) On the basis of RMSD evaluation, we decided to exclude mutants L3_35B and L3_35E, as they showed higher RMSD values in the last 5 ns of molecular dynamic simulations compared to the other structures. (b) RMSF comparison showed a significantly improved stability compared to the wild-type (WT) enzyme in all variants, although comparatively better for variants L3_35 A, L3_35D, and L3_35F. RMSD, root mean square deviation; RMSF, root mean square fluctuation [Color figure can be viewed at wileyonlinelibrary.com]

candidates L3_35B and L3_35E (Figure 4a). RMSF comparison in the amino acids range 55–75 showed that all candidates presented a lower RMSF value than the WT enzyme (Figure 4b). However, the L3_35C and L3_35G variants showed higher RMSF values relative to the other variants and thus they were discarded.

From the analysis of the structural alignment of the final frames for the remaining candidates (L3_35A, L3_35D, and L3_35F) with the WT enzyme, we decided to proceed with the experimental production of candidate L3_35A. Candidate L3_35F was discarded because of its higher global-RMSD from the WT structure compared to L3_35A (1.2 and 1.0 Å, respectively), while mutant L3_35D was discarded because of its higher structural deviation from the WT structure in the loop portion S413-P418, compared to L3_35A (Figure S2). Since this loop carries a highly conserved catalytic residue, R414, we tracked the explored rotamers during the MD trajectories, showing that L3_35A mutant allows a smaller number of possible rotamer conformations for the R414 side chain compared to L3_35D (Figure S3). Overall, the L3_35A features deletion of five amino acids and 35 point mutations with respect to the 5T1E structure. A structural

representation and a complete list of all the 35 single point mutations is shown in Figure S4.

3.4 | Production and purification

On the basis of our computational design analysis, we proceeded with the production and characterization of the L3_35A design. We obtained pure and active enzyme with an average yield of 5 mg/L of bacterial culture. The absorption spectrum of the pure sample, which shows two major absorbance peaks, one at 368 and one at 454 nm (Figure S5), clearly indicates the presence of flavin as a prosthetic group.

3.5 | Biochemical characterization

Apparent steady-state parameters (K_m, k_{cat}, and k_{cat}/K_m) measured against the substrate fructosyl-lysine were consistent with those reported in the literature for WT FPOX with available X-ray structures (Table 1), showing that our extensive design did not impair the catalytic activity of the enzyme. A comparison of the K_m of our design (4.1 mM) with those reported for 16 other FAOX/FPOX and related mutants for ε-fructosyl-lysine (range 0.19–27 mM; Lin & Zheng, 2010) shows that our engineered enzyme performs similarly to other members of the same family.

Enzyme stability was measured by the continuous assay previously described (Rigoldi et al., 2018), by incubating the protein at 37°C and detecting activity at different time intervals. In these conditions, the enzyme is completely inactivated after 6 hr (Figure S6a). Concerning the pH-activity profile, to the best of our knowledge, no comprehensive reports of the pH profile are available for other FPOX enzymes. A comparison with our previous work on Amadoriase I (Rigoldi, Gautieri et al., 2016; Rigoldi et al., 2018) and with other available data in the literature (Hirokawa, Nakamura, & Kajiyama, 2004) shows that our L3_35A design features a pH profile that is similar to other enzymes of the same family, with a marked loss of activity in acidic environments (pH < 5; see Figure S6b).

3.6 | X-ray structure determination and refinement

We used X-ray crystallography to confirm our structural predictions for mutant L3_35A. We obtained the crystal structure of the designed mutant at 1.38 Å resolution. The L3_35A protein crystallized in the $P4_22_12$ space group with one protein monomer in the asymmetric unit. Data collection and final refinement statistics are shown in Table S2. The structure of the designed region, as it appears in the electron density map, is shown in Figure S7. The deletion of one turn of an α-helix and the shortening of the following loop by a five-amino acid removal led to a wider and shorter access tunnel relative to the WT enzyme. The L3_35A mutant features a bottleneck of 3.02 Å (2.24 Å in the 5T1E structure) and a length of 5.40 Å (8.80 Å in the 5T1E structure).

The overall fold of the mutant is highly superimposable with the predicted model (using the equilibrated structure at the end of the 50 ns MD simulation as a reference), with an RMSD of 1.79 Å for the main chain atoms between the crystal structure and the MD-equilibrated conformation (Figure 5a). In particular, focusing on the designed loop region, the RMSD of the backbone between the crystal structure and the predicted structure is 1.12 Å (Figure 5b). Considering the whole MD simulations, the RMSD between the crystal structure and the MD trajectory is 0.94 Å.

The overall fold of the L3_35A mutant is highly superimposable with PnFPOX structure, with an RMSD of 1.71 Å (Figure 6a). The major differences are found in the region of loop redesign, where the two structures can be superimposed up to residue G60 (D in the L3_35A mutant), then show different folds, to become superimposable again starting from residue Q70 (P65 in the L3_35A structure, the amino acid that caps the helix; Figure 6b). The redesigned region also affects the folding of the E101-A117 helix (in the 5T1E numbering). In the WT enzyme, this segment forms hydrophobic interactions with the native loop (in particular, the hydrophobic interface is made up of residues V61, L63, A113, L114, and A117). In the structure of the mutant, this interface is mutated to smaller or polar residues (e.g., A61, A63, and N108), disrupting the hydrophobic interactions.

TABLE 1 Comparison of the engineered enzyme with the structurally characterized wild-type FPOX enzymes

Enzyme name	PDB ID	Sequence length	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}·mM^{-1})	Specific activity (U/mg)	References
L3_35A		426	4.1 ± 0.6	0.33 ± 0.02	0.080 ± 0.007	0.42 ± 0.02	
PnFPOX	5T1E	431	12.2 ± 1.8			2.56 ± 0.29	[a]
Amadoriase I	4WCT	445	4.2 0.51 ± 0.19	1.1 21.55 ± 3.08	41.68 ± 16.48		[b,c]
EtFPOX	4RSL	437					[d]

Note: Data for enzymatic activity are referred to fructosyl-lysine as substrate.
Abbreviation: FPOX, fructosyl peptide oxidase.
[a]Kim, Ferri, Tsugawa, Mori, and Sode (2010).
[b]Capuano et al. (2007).
[c]Rigoldi et al. (2018).
[d]Xing, Gan, Jia, Gao, and Gong (2013).

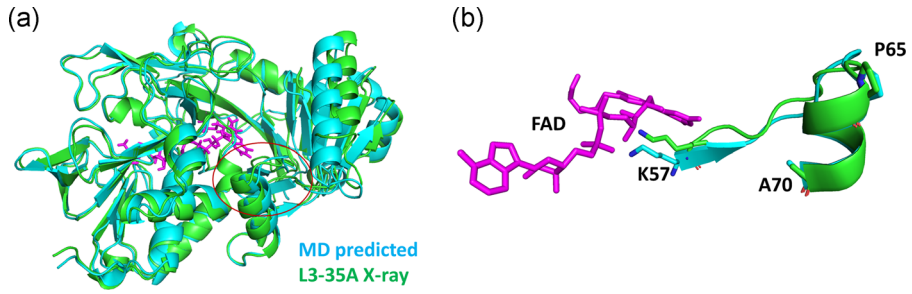

FIGURE 5 Superimposition of the crystal structure of L3_35A mutant (green cartoon) with its computationally predicted model (light-blue cartoon). The flavin cofactor is represented as purple sticks. (a) A comparison between the crystal structure of the L3_35A mutant (green cartoon) and the final frame of the 50-ns MD simulated structure (light-blue cartoon) shows a good agreement between the computationally predicted structure and the structure determined experimentally. For both structures, the flavin cofactor is completely superimposable. The red circled region is the modified loop highlighted in (b). (b) Zoom-in for the computationally designed region (residues K57 to A70) that was supposed to be the most affected conformation by our design approach (region terminal residues, A70, and K5 together with P65, used as helix capping, are represented as sticks). However, the superimposition of the two models reveals the reliability of our computational design approach. FAD, flavin adenine dinucleotide; MD, molecular dynamics [Color figure can be viewed at wileyonlinelibrary.com]

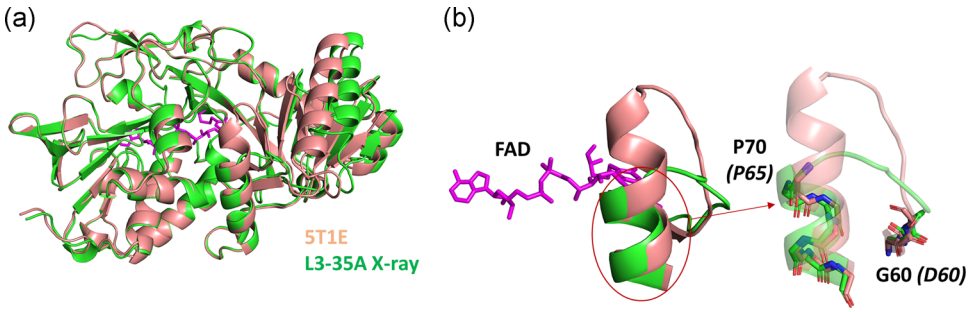

FIGURE 6 Superimposition of the crystal structure of L3_35A mutant (green cartoon) with the 5T1E crystal structure (pink-blue cartoon). The flavin cofactor is represented as purple sticks. The global superimposition of the two structures is shown in (a), providing clear evidence of their good structural homology. A detail of shorter helix and modeled loops is shown in (b). This region represents the highest structural deviated portion between the two structures: in particular, the structural difference affected only the remodeled loop and residues G60 to P70 (based on 5T1E numbering), corresponding to D60-P65 for l3-35A mutants. On the other hand, the helix, even if one-turn shorter in l3-35A, is highly aligned (backbone sticks). FAD, flavin adenine dinucleotide [Color figure can be viewed at wileyonlinelibrary.com]

We compared the catalytic residues, in terms of backbone fold and rotamer conformations, in the crystal structure of L3_35A mutant and in the 5T1E WT enzyme (Shimasaki et al., 2017). Since our goal is to preserve the specific activity of the WT enzyme in recognizing and cleaving glycated substrates, in our design approach, we applied structural constrains that would prevent rearrangement of both the backbone and the side-chains atoms of those residues that are involved in interactions with the sugar moiety of the substrate. The structural alignment shows no significant variation in the geometry of the four residues that interact with the sugar moiety and which are highly conserved among the members of the FPOX family (Shimasaki et al., 2017; residues: W235, E278, G372, and R415). The RMSD, restricted to these four catalytic residues, between the crystal structure of our mutant (L3_35A) and WT enzyme (5T1E) is under 1 Å (0.29 Å for backbone atoms and 0.67 Å considering also the heavy atoms of side chains), providing evidence that our design approach does not affect the key catalytic residues (Figure 7a). This result is supported by the high stability of these residues during the MD simulation at 300 K (the average RMSD during MD simulation is equal to 1.13 Å, excluding hydrogens).

Besides these key catalytic amino acids, there are other less conserved residues (viz., positions 257, 351, and 373) that have been reported to interact with the substrate and thus may play a role in the enzymatic activity (Shimasaki et al., 2017). For these positions, we observe an RMSD of 0.70 Å for backbone atoms (Figure 7b), suggesting that no change occurs in the local fold. Notably, L3_35A exhibits two important mutations with respect to the 5T1E structure, specifically A351G and H373D, while Y257 remains unchanged. Since these mutations do not impair the catalytic activity of our mutant, it is tempting to speculate that these positions may drive the selectivity over the specific glycated substrate (e.g., fructosyl valine or fructosyl-lysine) without playing a catalytic role.

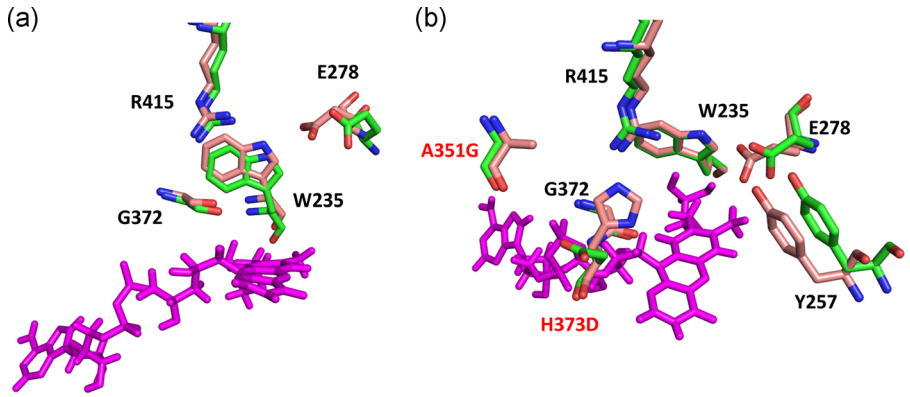

FIGURE 7 Superimposition of the highly conserved and mutable catalytic residues for (L3_35A mutant and 5T1E X-rays structures. (a) The superimposition of the four catalytic residues that are reported as key actors in the binding of the sugar moiety of the substrate (W235, E278, G372, and R415). The conformation of both the backbone and the rotamers of these residues is highly conserved between the wild-type enzyme (pink sticks) and our mutant (green sticks). The flavin cofactor, represented as purple sticks, is shown as a reference for the catalytic pocket position. (b) The design process introduced mutations in position 351, and 373 moving from Ala and His to Gly and Asp, respectively. Although not conserved, these residues have been suggested to play a role in defining the enzymatic activity. We note that despite the mutation of the side chains, the backbone conformations for both mutated residues are maintained [Color figure can be viewed at wileyonlinelibrary.com]

4 | CONCLUSION

Using a rational in silico computational design and screening method, we produced an engineered FPOX featuring a wider catalytic site access tunnel without compromising its global fold and catalytic function. The design involved the remodeling of the backbone structure of the enzyme, a feature that is not possible with conventional enzyme engineering techniques such as single point mutagenesis, and highly unlikely to occur using a directed evolution approach.

The L3_35A design proposed here represents a possible starting point for the development of improved biosensors for the monitoring of diabetes as well as for the design of therapeutic tools for the prevention or reduction of protein glycation in biological tissues. In future works, we plan to further engineer this design to improve the stability and to verify its ability to work on fructosyl dipeptides, fructosyl hexapeptides and, eventually, on the whole HbA1c and glycated albumin.

On the other hand, the proposed design strategy to modify the access tunnel appears to be a promising tool to expedite the design of novel and improved enzymes, with a significant reduction in cost and time for the experimental production and characterization of candidate enzyme variants. Our work lays in the general efforts to provide in silico strategies for the fast, accurate, and inexpensive development of improved enzymes that fit the needs of the industry. The ultimate goal is to increase the use of biocatalysts to improve chemical processes while working in compliance with the concepts and principles of Green Chemistry.

ACKNOWLEDGMENTS

This study was supported by Istituto Italiano di Tecnologia and Fondazione Cariplo (Grant 2016-0481). The final crystallographic coordinates of the crystal structures shown in the study are available in the RCSB PDB (accession code: 6Y4J).

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the RCSB Protein Data Bank (https://www.rcsb.org).

ORCID

Alfonso Gautieri http://orcid.org/0000-0003-0492-0130

REFERENCES

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., … Zwart, P. H. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D-Biological Crystallography*, *66*, 213–221.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., & Leslie, A. G. W. (2011). iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D-Biological Crystallography*, *67*, 271–281.

Capuano, E., Fedele, F., Mennella, C., Visciano, M., Napolitano, A., Lanzuise, S., … Fogliano, V. (2007). Studies on the effect of amadoriase from *Aspergillus fumigatus* on peptide and protein glycation in vitro. *Journal of Agricultural and Food Chemistry*, *55*, 4189–4195.

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., … Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, *26*, 1668–1688.

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., … Damborsky, J. (2012). CAVER 3.0: A tool for the analysis of transport pathways in dynamic protein structures. *PLoS Computational Biology*, *8*, 1002708.

Collard, F., Zhang, J., Nemet, I., Qanungo, K. R., Monnier, V. M., & Yee, V. C. (2008). Crystal structure of the deglycating enzyme fructosamine

oxidase (Amadoriase II). *Journal of Biological Chemistry*, *283*, 27007–27016.

DeLano, W. (2002). Pymol: An open-source molecular graphics tool. *Newsletter on protein crystallography*.

Del Turco, S., & Basta, G. (2012). An update on advanced glycation endproducts and atherosclerosis. *Biofactors*, *38*, 266–274.

Ferri, S., Kim, S., Tsugawa, W., & Sode, K. (2009). Review of fructosyl amino acid oxidase engineering research: A glimpse into the future of hemoglobin A1c biosensing. *Journal of Diabetes Science and Technology*, *3*, 585–592.

Ferri, S., Miyamoto, Y., Sakaguchi-Mikami, A., Tsugawa, W., & Sode, K. (2013). Engineering fructosyl peptide oxidase to improve activity toward the fructosyl hexapeptide standard for HbA1c measurement. *Molecular Biotechnology*, *54*, 939–943.

Gan, W., Gao, F., Xing, K., Jia, M., Liu, H., & Gong, W. (2015). Structural basis of the substrate specificity of the FPOD/FAOD family revealed by fructosyl peptide oxidase from eupenicillium terrenum. *Acta Crystallographica Section F: Structural Biology Communications*, *71*, 381–387.

Gautieri, A., Ionita, M., Silvestri, D., Votta, E., Vesentini, S., Fiore, G. B., … Redaelli, A. (2010). Computer-aided molecular modeling and experimental validation of water permeability properties in biosynthetic materials. *Journal of Computational and Theoretical Nanoscience*, *7*, 1287–1293.

Gautieri, A., Redaelli, A., Buehler, M. J., & Vesentini, S. (2013). Age- and diabetes-related nonenzymatic crosslinks in collagen fibrils: Candidate amino acids involved in advanced glycation end-products. *Matrix Biology*, *34*, 89–95.

Gautieri, A., Vesentini, S., Redaelli, A., & Buehler, M. J. (2009). Intermolecular slip mechanism in tropocollagen nanofibrils. *International Journal of Materials Research*, *100*, 921–925.

Goh, S. Y., & Cooper, M. E. (2008). The role of advanced glycation end products in progression and complications of diabetes. *Journal of Clinical Endocrinology and Metabolism*, *93*, 1143–1152.

Goldenzweig, A., Goldsmith, M., Hill, S. E., Gertman, O., Laurino, P., Ashani, Y., … Fleishman, S. J. (2016). Automated structure- and sequence-based design of proteins for high bacterial expression and stability. *Molecular Cell*, *63*, 337–346.

Hammond, J. B., & Kruger, N. J. (1988). The Bradford method for protein quantitation. *Methods in Molecular Biology*, *3*, 25–32.

Harvey, M. J., Giupponi, G., & De Fabritiis, G. (2009). ACEMD: Accelerating biomolecular dynamics in the microsecond time scale. *Journal of Chemical Theory and Computation*, *5*, 1632–1639.

Hirokawa, K., Nakamura, K., & Kajiyama, N. (2004). Enzymes used for the determination of HbA1C. *FEMS Microbiology Letters*, *235*, 157–62.

Huang, P. S., Ban, Y. E. A., Richter, F., Andre, I., Vernon, R., Schief, W. R., & Baker, D. (2011). RosettaRemodel: A generalized framework for flexible backbone protein design. *PLoS One*, *6*, e24109.

Kaufmann, K. W., Lemmon, G. H., Deluca, S. L., Sheehan, J. H., & Meiler, J. (2010). Practically useful: What the Rosetta protein modeling suite can do for you. *Biochemistry*, *49*, 2987–98.

Kim, S., Ferri, S., Tsugawa, W., Mori, K., & Sode, K. (2010). Motif-based search for a novel fructosyl peptide oxidase from genome databases. *Biotechnology and Bioengineering*, *106*, 358–366.

Kim, S., Miura, S., Ferri, S., Tsugawa, W., & Sode, K. (2009). Cumulative effect of amino acid substitution for the development of fructosyl valine-specific fructosyl amine oxidase. *Enzyme and Microbial Technology*, *44*, 52–56.

Kokkonen, P., Bednar, D., Pinto, G., Prokop, Z., & Damborsky, J. (2019). Engineering enzyme access tunnels. *Biotechnology Advances*, *37*, 107386.

Lin, Z. L., & Zheng, J. (2010). Occurrence, characteristics, and applications of fructosyl amine oxidases (amadoriases). *Applied Microbiology and Biotechnology*, *86*, 1613–1619.

Mandell, D. J., Coutsias, E. A., & Kortemme, T. (2009). Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nature Methods*, *6*, 551–552.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, *40*, 658–674.

Miura, S., Ferri, S., Tsugawa, W., Kim, S., & Sode, K. (2006). Active site analysis of fructosyl amine oxidase using homology modeling and site-directed mutagenesis. *Biotechnology Letters*, *28*, 1895–1900.

Miura, S., Ferri, S., Tsugawa, W., Kim, S., & Sode, K. (2008). Development of fructosyl amine oxidase specific to fructosyl valine by site-directed mutagenesis. *Protein Engineering, Design & Selection*, *21*, 233–239.

Monnier, V. M., & Wu, X. (2003). Enzymatic deglycation with amadoriase enzymes from Aspergillus sp. as a potential strategy against the complications of diabetes and aging. *Biochemical Society Transactions*, *31*, 1349–1353.

Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., … Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Biological Crystallography*, *67*, 355–367.

Nagaraj, R. H., Linetsky, M., & Stitt, A. W. (2012). The pathogenic role of Maillard reaction in the aging eye. *Amino Acids*, *42*, 1205–1220.

Nelson, M. T., Humphrey, W., Gursoy, A., Dalke, A., Kale, L. V., Skeel, R. D., & Schulten, K. (1996). NAMD: A parallel, object oriented molecular dynamics program. *International Journal of High Performance Computing Applications*, *10*, 251–268.

Ogawa, N., Kimura, T., Umehara, F., Katayama, Y., Nagai, G., Suzuki, K., … Ichimura, M. (2019). Creation of haemoglobin A1c direct oxidase from fructosyl peptide oxidase by combined structure-based site specific mutagenesis and random mutagenesis. *Scientific Reports*, *9*, 942.

Paul, R. G., & Bailey, A. J. (1996). Glycation of collagen: The basis of its central role in the late complications of ageing and diabetes. *International Journal of Biochemistry and Cell Biology*, *28*, 1297–1310.

Qian, Y., Zheng, J., & Lin, Z. L. (2013). Loop engineering of Amadoriase II and mutational cooperativity. *Applied Microbiology and Biotechnology*, *97*, 8599–8607.

Rigoldi, F., Donini, S., Giacomina, F., Sorana, F., Redaelli, A., Bandiera, T., … Gautieri, A. (2018). Thermal stabilization of the deglycating enzyme Amadoriase I by rational design. *Scientific Reports*, *8*, 3042. https://doi.org/10.1038/s41598-018-19991-x

Rigoldi, F., Gautieri, A., Dalle Vedove, A., Lucarelli, A. P., Vesentini, S., & Parisini, E. (2016). Crystal structure of the deglycating enzyme Amadoriase I in its free form and substrate-bound complex. *Proteins Structure Function and Bioinformatics*, *84*, 744–758.

Rigoldi, F., Spero, L., Dalle Vedove, A., Redaelli, A., Parisini, E., & Gautieri, A. (2016). Molecular dynamics simulations provide insights into the substrate specificity of FAOX family members. *Molecular BioSystems*, *12*, 2622–2633.

Schrödinger. (2017). *Pymol*. Retrieved from https://pymol.org/2

Sell, D. R., & Monnier, V. M. (2012). Molecular basis of arterial stiffening: Role of glycation—A mini-review. *Gerontology*, *58*, 227–237.

Shahbazmohammadi, H., Sardari, S., Lari, A., & Omidinia, E. (2019). Engineering an efficient mutant of Eupenicillium terrenum fructosyl peptide oxidase for the specific determination of hemoglobin A1c. *Applied Microbiology and Biotechnology*, *103*, 1725–1735.

Shimasaki, T., Yoshida, H., Kamitori, S., & Sode, K. (2017). X-ray structures of fructosyl peptide oxidases revealing residues responsible for gating oxygen access in the oxidative half reaction. *Scientific Reports*, *7*, 2790.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. Z., … Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Molecular Systems Biology*, *7*, 539.

Snedeker, J. G., & Gautieri, A. (2014). The role of collagen crosslinks in ageing and diabetes—The good, the bad, and the ugly. *Muscles, Ligaments and Tendons Journal*, *4*, 303–308.

Takahashi, M., Pischetsrieder, M., & Monnier, V. M. (1997). Isolation, purification, and characterization of amadoriase isoenzymes (fructosyl

amine-oxygen oxidoreductase EC 1.5.3) from *Aspergillus* sp. *Journal of Biological Chemistry*, *272*, 3437–3443.

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., … Adams, P. D. (2007). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica Section D: Biological Crystallography*, *D64*, 61–69.

Vincent, A. M., Russell, J. W., Low, P., & Feldman, E. L. (2004). Oxidative stress in the pathogenesis of diabetic neuropathy. *Endocrine Reviews*, *25*, 612–628.

Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, *25*, 1157–1174.

Wu, X. L., Palfey, B. A., Mossine, V. V., & Monnier, V. M. (2001). Kinetic studies, mechanism, and substrate specificity of Amadoriase I from *Aspergillus* sp. *Biochemistry*, *40*, 12886–12895.

Xing, K., Gan, W., Jia, M., Gao, F., & Gong, W. (2013). Expression, purification, crystallization and preliminary X-ray diffraction analysis of EtFPOX from *Eupenicillium terrenum* sp. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, *69*, 666–668.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Rigoldi F, Donini S, Torretta A, et al. Rational backbone redesign of a fructosyl peptide oxidase to widen its active site access tunnel. *Biotechnology and Bioengineering*. 2020;1–11. https://doi.org/10.1002/bit.27535