Sortase mutants with improved protein thermostability and enzymatic activity obtained by consensus design

Wójcik, Magdalena; Vázquez Torres, Susana; Quax, Wim J; Boersma, Ykelien L

Published in:
Protein Engineering, Design & Selection

DOI:
10.1093/protein/gzaa018

Publication date:
2020

peds
protein engineering design & selection

Protein Engineering, Design and Selection, 2019, vol. 32 no. 12, pp. 555–564
doi: 10.1093/protein/gzaa018
Original Article

OXFORD

Original Article

Sortase mutants with improved protein thermostability and enzymatic activity obtained by consensus design

Magdalena Wójcik, Susana Vázquez Torres, Wim J. Quax, and Ykelien L. Boersma*

Department of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, A. Deusinglaan 1, 9713 Groningen, The Netherlands

*To whom correspondence should be addressed. E-mail: y.l.boersma@rug.nl

Received 3 November 2019; Revised 16 March 2020; Editorial Decision 3 July 2020; Accepted 3 July 2020

Abstract

Staphylococcus aureus sortase A (SaSrtA) is an enzyme that anchors proteins to the cell surface of Gram-positive bacteria. During the transpeptidation reaction performed by SaSrtA, proteins containing an N-terminal glycine can be covalently linked to another protein with a C-terminal LPXTG motif (X being any amino acid). Since the sortase reaction can be performed *in vitro* as well, it has found many applications in biotechnology. Although sortase-mediated ligation has many advantages, SaSrtA is limited by its low enzymatic activity and dependence on Ca^{2+}. In our study, we evaluated the thermodynamic stability of the SaSrtA wild type and found the enzyme to be stable. We applied consensus analysis to further improve the enzyme's stability while at the same time enhancing the enzyme's activity. As a result, we found thermodynamically improved, more active and Ca^{2+}-independent mutants. We envision that these new variants can be applied in conjugation reactions in low Ca^{2+} environments.

Key words: Conjugation, consensus design, protein engineering, protein stability, *Staphylococcus aureus* sortase A

Introduction

The bacterial transpeptidase sortase A (SrtA) enzyme is found in the membrane of most Gram-positive bacteria, where it mediates the anchoring of proteins and virulence factors to the cell wall. The best characterized and most studied SrtA is that of *Staphylococcus aureus* (SaSrtA) (Jacobitz *et al.*, 2017; Mazmanian *et al.*, 1999). This enzyme recognizes a C-terminal sortase-specific motif, which consists of the five amino acids LPXTG (X being any amino acid), and cleaves the peptide bond between the penultimate (Thr) and last (Gly) amino acid. After cleavage, the substrate remains bound to the enzyme until this intermediate is resolved by a nucleophilic attack from a second substrate, containing an N-terminal pentaglycine motif (Marraffini *et al.*, 2006; Suree *et al.*, 2009). Interestingly, the sortase-mediated reaction can be used *in vitro* for site-specific conjugation reactions (Call and Klaenhammer, 2013; Proft, 2009); however, the soluble version of $SaSrtA_{\Delta N59}$ wild type (WT) lacking the first 59 residues is a catalytically inefficient enzyme (Chen *et al.*, 2011; Chen *et al.*, 2016; Jeong *et al.*, 2017; Wu *et al.*, 2017a).

Both directed evolution and rational design have been employed to optimize certain enzyme characteristics (Arnold and Volkov, 1999; Chen, 2001; Cobb *et al.*, 2013; Steiner and Schwab, 2012). The choice of strategy depends on the availability of a robust high-throughput screening or selection method and of data on the enzyme's structure-function relationship; naturally, both methods can be used simultaneously in a semi-rational approach (Lutz, 2011). So far, $SaSrtA_{\Delta N59}$ has been successfully engineered to create mutants with an altered substrate specificity(An *et al.*, 2014) as well as catalytically more active mutants using directed evolution (Antos *et al.*, 2016; Chen *et al.*, 2011; Chen *et al.*, 2016). An evolved $SaSrtA_{\Delta N59}$ pentamutant (PM) was reported to have a 120-fold higher $k_{cat}/K_{m\ LPETG}$ and a 20-fold higher $K_{m\ GGG}$ (Chen *et al.*, 2011). However, a higher enzymatic activity sometimes comes at

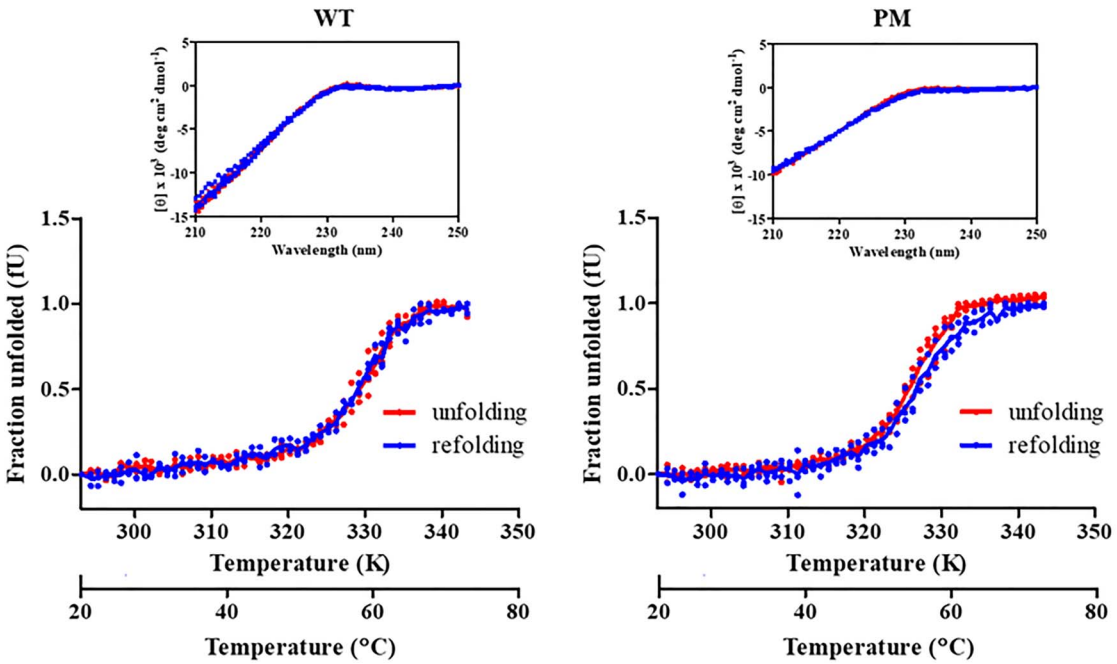

Fig. 1 Reversibility of protein unfolding measured by CD. Individual graphs represent the unfolding of native (red line) and unfolding of refolded (blue line) WT (left) and PM (right) (Romero-Romero *et al.*, 2015). The inset figures show the CD spectra of the native (red) and refolded proteins (blue).

the expense of stability (Steiner, 2012), which can have implications for reuse, immobilization or storage (Xu *et al.*, 2015). The stability of SaSrtA$_{\Delta N59}$ WT was previously characterized using circular dichroism (Bentley *et al.*, 2008; Frankel *et al.*, 2007). We also used far-ultraviolet circular dichroism (CD) (Li and Hirst, 2017; Tripathi, 2013) and determined the reversibility of thermal unfolding of both truncated SaSrtA$_{\Delta N59}$ WT and SaSrtA$_{\Delta N59}$ PM subjected to increasing temperatures. We found that both enzymes are thermodynamically stable proteins, with a reversible change in CD as a function of temperature.

Previous efforts to improve the melting temperature (T_m) include linking two intradomain cysteine residues using biselectrophiles (Pelay-Gimeno *et al.*, 2018). Here, we applied consensus-based mutagenesis for the improvement of the stability of SaSrtA$_{\Delta N59}$ WT: we generated multiple sequence alignments (MSA) to assess the effect of consensus mutations on the stability and activity of SaSrtA$_{\Delta N59}$ (Wu *et al.*, 2017b). We attempted to further improve the stability of both WT and PM enzymes while at the same time improving the enzyme's activity. We applied consensus design (Lehmann and Wyss, 2001; Porebski and Buckle, 2016) to search for evolutionary information preserved in homologous sequences (Jäckel *et al.*, 2010). Therefore, we built SrtA-MSAs consisting of a panel of homologous SrtAs derived from organisms adapted to different temperatures, such as mesophiles and thermophiles. Based on our MSAs, we selected and replaced eight non-consensus amino acids of SaSrtA$_{\Delta N59}$ WT with the corresponding consensus ones. As a result, we found some mutations which not only led to an improvement of the thermostability of the protein but also to an improvement of the enzymatic activity (Porebski and Buckle, 2016; Steipe *et al.*, 1994). Finally, we introduced the two best mutations into the known SaSrtA$_{\Delta N59}$ PM, making it even more active than previously reported (Chen *et al.*, 2011) and less dependent on Ca^{2+} ions (Jeong *et al.*, 2017).

Materials and Methods

Consensus design

Sequences to be used in the MSA encoding class A sortases were retrieved from the UniProt database (UniProt Consortium, 2009). All sequences were verified for their proper distribution of soluble and insoluble domains using the TMHMM Server v. 2.0 (Krogh *et al.*, 2001). This server uses a hidden Markov model for the prediction of integral membrane proteins. As a result, 134 sequences encoding SrtA from different Gram-positive organisms were selected, with 35 enzymes from extremophile organisms. All organisms from which SrtA sequences were selected and used in the MSA are shown in Fig. 2.

Prior to generation of the MSA using the PROMALS3D tool (Pei *et al.*, 2008a), secondary structure elements were aligned, and highly conserved residues in SrtA were identified using high-resolution 3D structures. To this purpose, *S. aureus* SrtA (PDB 1T2P), *Bacillus anthracis* SrtA (PDB 2KW8), *Streptococcus agalactiae* SrtA (PDB 3RCC), *Streptococcus pneumoniae* SrtA (PDB 4O8L), *Streptococcus mutans* SrtA (PDB 4TQX), *Listeria monocytogenes* SrtA (PDB 5HU4), *Corynebacterium diphtheriae* SrtA (PDB 5K9A) and *Actinomyces oris* SrtA (PDB 5UTT) were selected. Based on predictions obtained from the structural alignment of these eight enzymes, all 134 retrieved sequences were used for the construction of the MSA.

Cloning, production and purification of SaSrtA$_{\Delta N59}$ WT, SaSrtA$_{\Delta N59}$ PM and selected mutants

The DNA sequence encoding SaSrtA$_{\Delta N59}$ WT (Ilangovan *et al.*, 2001) was cloned into the pET28a plasmid (Novagen, USA) between the *Bam*HI and *Nde*I restriction sites; the previously described mutations of SaSrtA$_{\Delta N59}$ PM (P94R/D160N/D165A/K190E/K196T) and Ca^{2+}-independent SaSrtA$_{\Delta N59}$ mutant (E105K/E108Q, indicated as

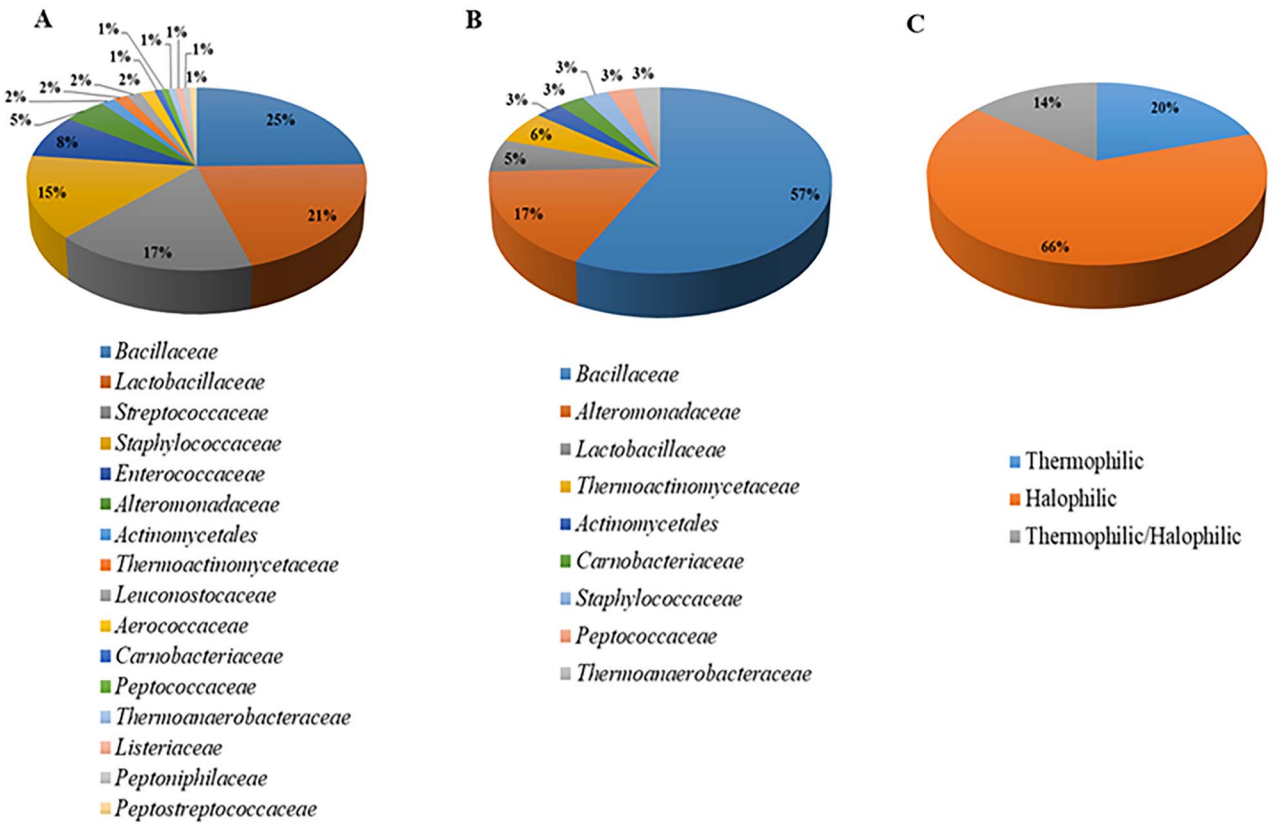

Fig. 2 Organisms from which SrtA sequences were selected and used in the MSA. (A) Pie chart representing the 17 different taxonomic families of the 134 bacterial SrtA sequences selected from the UniProt database and analyzed using the TMHMM Server. (B) Pie chart representing the classification of 35 extremophile SrtA sequences. (C) Pie chart representing the classification of 35 bacterial SrtA sequences according to the extremophile group they belong to.

WT^{-Ca}) were introduced via Quikchange® site-directed mutagenesis (Agilent, La Jolla, CA, USA) (Liu and Naismith, 2008). After confirmation of the sequence, all plasmids were used for the transformation of competent *Escherichia coli* BL21(DE3) cells (New England Biolabs, USA). The $SaSrtA_{\Delta N59}$ mutations selected from the MSA analysis were created via QuikChange® site-directed mutagenesis using templates encoding truncated $SaSrtA_{\Delta N59}$ WT or $SaSrtA_{\Delta N59}$ PM. The primers were designed according to the QuikChange® instruction manual, and their sequences can be found in the supporting information (Table S1). After induction with 1-mM isopropyl β-D-1-thiogalactopyranoside (IPTG), protein production was continued for 4 h at 37°C. Cells were then lysed by sonication, and the cleared lysate was purified using Ni-NTA resin (Qiagen, Hilden, Germany) according to the manufacturer's instructions. All proteins were purified to 95% purity via size exclusion chromatography using a Superdex75 HiLoad 16/60 column (GE Healthcare, MA, USA) in Tris-HCl buffer (50 mM, pH 7.5) supplemented with 150 mM NaCl and 10% v/v glycerol. Protein concentrations were determined using a Pierce bicinchoninic acid protein assay kit (Thermo Scientific).

Thermal stability

Thermal stability of all proteins was determined using a Jasco J-815 far-UV CD Spectropolarimeter equipped with a temperature controller, a thermostatted cell holder and a thermostatic bath. The instrument was under constant nitrogen flush. Experiments were performed in quartz cuvettes with a 1-mm path length. The protein concentration of WT $SaSrtA_{\Delta N59}$ and the mutants were set to 1 mg/mL in a 50 mM Tris-HCl buffer, pH 7.5, supplemented with 150 mM NaCl. The secondary structure of $SaSrtA_{\Delta N59}$ was recorded at 25°C as a spectrum between 210 and 250 nm; each measurement was repeated three times and corrected for solvent contributions. The molar ellipticity (θ) was calculated as reported before (Greenfield, 2006).

The reversibility of the unfolding process was assessed after subjecting the proteins to a temperature increase from 20 to 70°C with a ramp of 1°C/min and cooling. Subsequently, thermal unfolding was monitored at the minimum measured at 210 nm, which corresponds to secondary structure content. Once spectra from the initial and the new measurement at 210 nm matched, thermodynamics of the samples could be calculated. Since spectral changes measured for $SaSrtA_{\Delta N59}$ showed sigmoidal curves with a single transition from a native to an unfolded state, we assumed a two-state model for protein denaturation. The equilibrium constant of unfolding (K_{eq}) was calculated from the fraction of folded and unfolded protein (Klotz, 1996; Liu *et al.*, 2008; Wong *et al.*, 2004). The enthalpy ΔH_u for a globular monomeric protein was calculated using the Van't Hoff plot (supporting information, Fig. S1), and the Gibbs free energy of unfolding ΔG_u was estimated using the Gibbs-Helmholtz equation at 25°C. For calculation of the change in heat capacity upon unfolding ΔC_p, the values of the coefficients were obtained through the relationship between reported parametric equations (Milardi *et al.*, 1997) (supporting information, Table S2).

Position in the alignment	164	198	206	219	235	248	284	297	302
Amino acid residue in *SaSrtA* $_{\Delta 59}$	**91**	**113**	**121**	**129**	**144**	**155**	**184**	**193**	**198**
Pediococcus acidilactici	V	G	H	L	L	V	C	R	G
Lactobacillus helveticus	M	G	Y	L	L	I	C	R	G
Bacillus licheniformis	T	G	H	L	I	V	C	R	G
Marinactinospora thermotolerans	T	G	R	I	V	V	C	R	A
Thermoactinomyces sp. DSM 45891	A	G	R	K	I	V	C	R	S
Sediminibacillus halophilus	T	G	R	M	L	I	C	R	A
Halobacillus dabanensis	A	R	R	V	Y	V	C	R	S
Virgibacillus salinus	T	D	R	V	L	I	C	R	A
Halobacillus aidingensis	T	R	R	V	Y	V	C	R	S
Thalassobacillus cyri	T	D	R	V	V	V	C	R	A
Terribacillus saccharophilus	T	D	R	F	L	V	C	R	G
Marinilactibacillus piezotolerans	L	G	H	L	V	I	C	R	A
Gracilibacillus orientalis	T	D	R	V	V	V	C	R	A
Gracilibacillus ureilyticus	T	G	R	V	F	A	C	R	T
Amphibacillus marinus	T	D	R	H	L	V	C	R	G
Halobacillus alkaliphilus	A	R	R	V	V	V	C	R	S
Thermoflavimicrobim dichotomicum	T	G	R	L	I	I	C	R	A
Halolactibacillus miurensis	T	S	R	V	V	M	C	R	A
Gracilibacillus kekensis	T	D	R	V	I	M	C	R	A
Ornithinibacillus halophilus	T	D	R	V	V	I	C	R	A
Salinicoccus alkaliphilus DSM 16010	P	Q	R	R	L	I	C	R	A
Desulfitobacterium chlororespirans DSM 11544	V	G	R	Q	V	V	C	R	G
Tepidanaerobacter syntrophicus	A	G	R	Q	I	V	C	R	A
Salimicrobium flavidum	T	G	R	F	L	V	C	R	G
Aestuariibacter aggregatus	S	G	R	H	I	I	C	R	A
Marinobacter zhejiangensis	V	G	R	H	L	V	C	R	A
Marinobacter mobilis	V	G	R	H	L	V	C	R	A
Marinobacter pelagius	T	G	R	H	L	V	C	R	A
Marinobacter gudaonensis	T	G	R	H	L	V	C	R	A
Natribacillus halophilus	T	K	R	V	V	V	C	R	G
Piscibacillus halophilus	A	G	R	L	I	V	C	R	A
Terribacillus halophilus	T	D	R	F	L	V	C	R	G
Oceanobacillus limi	T	D	R	V	V	I	C	R	A
Lentibacillus halodurans	T	D	R	V	V	I	C	R	A
Marinobacterium lutimaris	D	A	R	H	L	V	C	R	A
Staphylococcus aureus	P	Q	T	Q	F	M	C	V	K
Consensus	**h**	**s**	**+**	**h**	**l**	**l**	**C**	**R**	**t**
Conservation index	**<5**	**<5**	**5**	**<5**	**<5**	**5**	**9**	**8**	**5**

Fig. 3 Part of the consensus design analysis of strategy 2, showing the positions selected for the mutagenesis in the SaSrtA$_{\Delta N59}$ WT shown in green, with orange indicating the residues in the SaSrtA$_{\Delta N59}$ WT. The catalytic Cys is highlighted in red with Cys184 residue in the SaSrtA$_{\Delta N59}$ WT depicted in yellow. For the consensus analysis, bold and uppercase letters represent conserved amino acid residues: 'l' aliphatic residues (I, V, L), 'h' hydrophobic residues (W, F, Y, M, L, I, V, A, C, T, H), 's' small residues (A, G, C, S, V, N, D, T, P), '+' positively charged residues (K, R, H), 't' tiny residues (A, G, C, S).

Library screening and fluorescence-based activity assay

The enzymatic activity of purified SaSrtA$_{\Delta N59}$ WT and mutants was measured using a fluorometric assay as described previously (Ton-That *et al.*, 2000; Wójcik *et al.*, 2019). Briefly, the cleavage of the internally quenched substrate Abz-LPETG-Dap(Dnp) (Bachem AG, Switzerland) (20 μM) was monitored using a FLUOstar Omega Spectrometer (BMG LABTECH), measuring the change in fluorescence at λ_{em} = 460 nm after excitation at λ_{ex} = 355 nm at 37 or 45°C. The NH_2-5Gly-OH (Bachem AG, Switzerland) nucleophile was added to the reaction to a final concentration of 2 mM. The purified enzymes were individually added to a final concentration of 1 μM. Fluorescence signals were plotted against time, and the slope values were obtained from the linear phase of the reaction.

To screen the mutant libraries, the activity of individual SrtA mutants was assessed in cell lysate. Briefly, *E. coli* BL21(DE3) transformants were grown overnight at 37°C in a 96-deep well Masterblock® (Greiner Bio-One), after which the overnight cultures were diluted 1:100 in 1-mL 2 YT media supplemented with Kanamycin (30 μg/mL) and incubated for 45 min at 37°C in an orbital shaker. Protein production was induced by the addition of IPTG to a final concentration of 0.5 mM and continued for 4 h. After centrifugation, pellets were resuspended in 50 μL of BugBuster® (Novagen) supplemented with 1 mM EDTA and incubated for 30 min at room temperature with orbital shaking. Finally, 950 μL activity assay buffer with 5 mM $CaCl_2$ was added to each well and the Masterblock® was centrifuged for 30 min at a speed of 3000 rpm. The activity of the SrtA variants was measured using a fluorescence-based activity assay in 96-well black polystyrene plates (Greiner). Each well was filled with 50 μL of the cell lysate, assay buffer and internally quenched Abz-LPETG-Dap(Dnp) substrate to a final concentration of 20 μM. The expression level of the proteins was

Table I. Non-consensus residues in $SaSrtA_{\Delta N59}$ WT identified by both strategies

Position	Type of residue in the consensus sequence (Strategy 1/Strategy 2)	Modifications
P91	Hydrophobic/hydrophobic	All possible amino acids
Q113	Hydrophobic/small	All possible amino acids
Q129	Hydrophobic/hydrophobic	All possible amino acids
F144	Aliphatic/aliphatic	I, V, L
M155	Aliphatic/aliphatic	I, V, L
K198	Hydrophobic/tiny	All possible amino acids

Hydrophobic residues are W, F, Y, M, L, I, V, A, C, T and H; small residues are A, G, C, S, V, N, D, T and P; aliphatic residues are I, V and L; tiny residues are A, G, C and S.

Consensus amino acids at positions 144 and 155 showed more specific preferences for amino acid substitutions. The other residues listed in this table were modified into all 20 natural amino acids.

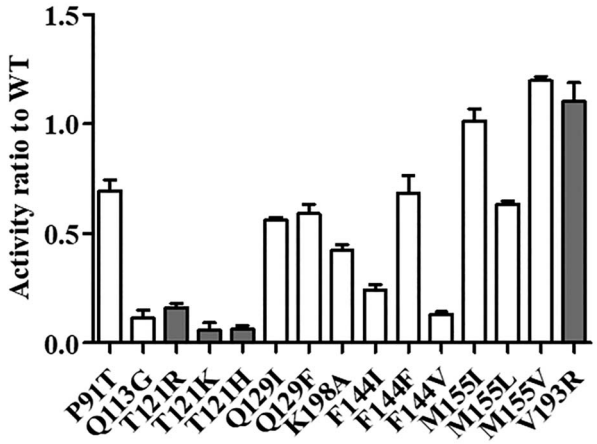

Fig. 4 Bar graph showing the residual activity of purified wild type (WT) and mutants selected from the MSAs ($n = 3$). The white bars represent variants emerging from both strategies 1 and 2. The gray bars represent variants emerging from strategy 2. Variants M155V and V193R, showing higher activity than the WT, were selected for further analysis.

monitored by SDS-PAGE, as shown in the supporting information (Fig. S2B). For the measurement of activity in the absence of Ca^{2+} ions, 5 mM $CaCl_2$ was exchanged for 5 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA). For the measurement of activity at 60°C, the enzymes were pre-incubated in a water bath at 60°C for 90 min.

Results

Thermostability of sortase WT and sortase PM

For characterization of the secondary structure content and thermostability of the purified $SaSrtA_{\Delta N59}$ WT and $SaSrtA_{\Delta N59}$ PM proteins, far UV-circular dichroism spectra were used. It showed minima at 210 nm which are characteristic for the presence of β-sheets. Additionally, the reversibility of unfolding was monitored by gradually bringing samples to a temperature of 70°C with steps of 1°C/min, then cooling down to 20°C and repeating the process; as shown in Fig. 1, the unfolding and refolding transitions were found to be superimposable, which indicates that the proteins exhibited reversible unfolding.

The $SaSrtA_{\Delta N59}$ WT was therefore corroborated to be a thermodynamically stable protein, refolding after being subjected to 70°C and cooled down to 25°C. The same behavior was observed for the $SaSrtA_{\Delta N59}$ PM.

Identification and selection of mutagenesis sites

Protein consensus design was performed to identify and select those amino acid positions in the structure of the $SaSrtA_{\Delta 59}$ WT, which could affect the thermostability of the enzyme. In the UniProt database, 166 SrtA sequences from different bacteria were found. All sequences were then analyzed using the TMHMM Server; only sequences with a proper distribution of cytoplasmic, transmembrane and extracellular domains were used for further studies. Consequently, a total of 134 sequences were selected for the MSA. The taxonomic classification of all organisms that carry the *srtA* gene is shown in Fig. 2.

The 134 selected sequences from different organisms were then classified according to the organism's habitat: extremophilic or mesophilic (Rampelotto, 2013). As shown in Fig. 2B, the extremophile organisms showed taxonomical diversity. Most of these organisms were halophiles, while some were both halophilic and thermophilic, and others solely thermophilic (Fig. 2C). Based on these categories, we developed two strategies: strategy 1 covered all 134 sequences encoding SrtA genes, while strategy 2 only covered the 35 SrtA sequences from extremophiles.

We used PROMALS3D to build the MSAs (Pei *et al.*, 2008b). This progressive tool can be used to accurately construct alignments using information from available 3D structures, database homologs and predicted secondary structures (Pei *et al.*, 2008a). For each amino acid, the conservation index, ranging from 0 to 9 with 9 indicating the highest degree of conservation, was calculated with the AL2CO program (Pei and Grishin, 2001) linked to the PROMALS3D server. We used the conservation index to assess the quality of our alignment and to select positions for the mutagenesis. The conservation index was used together with the consensus analysis to select amino acid substitutions (Fig. 3).

As expected, the conservation index for the Cys residue, which is the catalytic center of all SrtA enzymes, was the highest. Interestingly, the second highest conservation index after analyzing the *srtA* sequences of extremophile organisms was given to the amino acid residue in position 193, where all analyzed sequences except the $SaSrtA_{\Delta N59}$ WT contained Arg. Some positions selected in this study had a conservation index below 5, but here consensus outcome was used as a decisive variable (Table I). After analyzing the MSAs from both strategies described above, we identified 29 non-conserved positions in the input $SaSrtA_{\Delta N59}$ WT sequence (PDB structure 1T2P). More precisely, analysis of the 134 sequences selected for strategy 1 resulted in the identification of 13 positions with non-conserved residues in $SaSrtA_{\Delta N59}$ WT (supporting information Table S3). The advantage of this strategy was the considerable number of input sequences and high taxonomic diversity of the analyzed sequences. Analysis of the sequences selected for strategy 2 led to the identification of 16 positions at which $SaSrtA_{\Delta N59}$ WT preferred a different amino acid compared to the consensus (supporting information

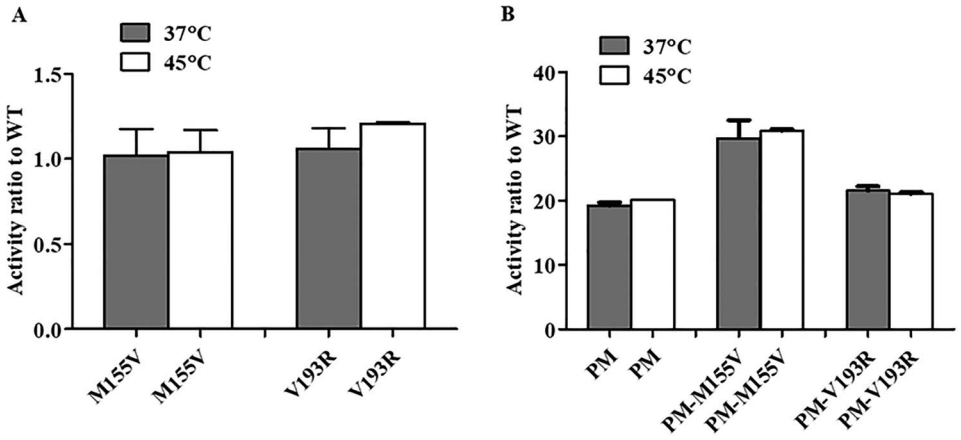

Fig. 5 Bar graphs showing the residual activity the mutants compared to the wild type (WT) ($n = 3$). (A) The enzymatic activity of single mutants M155V and V193R compared to WT, measured at 37°C (gray bars) and 45°C (white bars). (B) The enzymatic activity of the $SaSrtA_{\Delta N59}$ PM and PM with incorporated single mutations (PM-M155V and PM-V193R) compared to WT and measured at 37°C (gray bars) and 45°C (white bars).

Table S4). Closer examination of the MSA from strategy 2 showed that at position 121, all analyzed extremophile sequences showed a very strong preference for amino acids with positively charged side chains (Fig. 3). Therefore, we substituted the polar, uncharged Thr in $SaSrtA_{\Delta N59}$ WT with Arg, His and Lys. Of particular interest was position 193, since all analyzed sequences had an Arg residue at this position, whereas in the $SaSrtA_{\Delta N59}$ sequence, a hydrophobic Val was incorporated. We thus decided to introduce the specific point mutations T121R, T121K, T121H and V193R via Quikchange® mutagenesis instead of constructing libraries.

Next, we compared the results from both strategies and selected the six non-consensus residues in $SaSrtA_{\Delta N59}$ WT that emerged in the MSAs from both strategies for further studies (Table I). Since positions 91, 113, 129 and 198 showed a high variation in (types of) amino acids, site-saturation libraries were constructed on these positions. At the other two positions, 144 and 155, a more specific preference towards aliphatic amino acids was observed, and therefore, these residues were mutated into the three aliphatic amino acids Ile, Val and Leu.

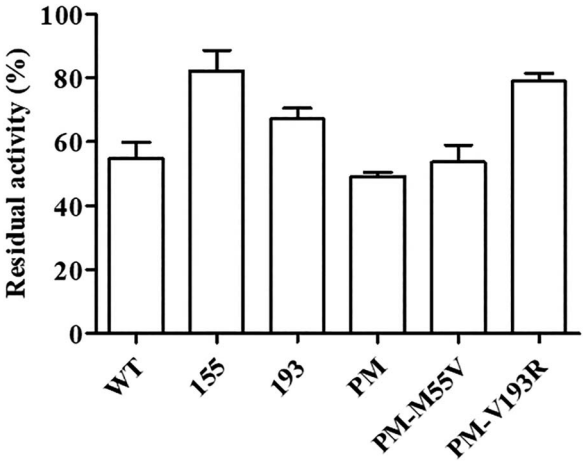

Fig. 6 Bar graph representing the residual enzymatic activity of the mutants after 90 min incubation at 60°C ($n = 3$). The activity of enzymes before the temperature treatment was set as 100% ($t = 0$ min).

Enzymatic activity of the mutants

After purification, $SaSrtA_{\Delta N59}$ WT and single mutants (Table I) were first examined for their enzymatic activity on the quenched substrate. The three mutants made at position T121 were all found to be inactive. On the other hand, mutants M155V and V193R were found to be slightly more active than the WT (Fig. 4).

Next, the libraries at positions 91, 113, 129 and 198 were screened for their activity using cell lysate (supporting information Fig. S1). As a result, five potential hits, P91T, Q113G, Q129I, Q129F and K198A, were found, which were purified and assessed for their activity. Unfortunately, the purified mutants showed lower activities than the WT (Fig. 4). Finally, of the six mutants at positions F144 and M155, only mutant M155V was found to have an improved activity in comparison to the WT (Fig. 4).

Mutations M155V and V193R were then incorporated in the previously described $SaSrtA_{\Delta N59}$ PM, and the activity of the resulting mutants PM-M155V and PM-V193R was assessed at 37°C and 45°C. As shown in Fig. 5, the temperature change did not have a significant influence on the initial rate of the reaction. Interestingly, variant PM-M155V performed around 50% better than $SaSrtA_{\Delta N59}$ PM at either temperature, while the activity of variant PM-V193R was comparable to that of $SaSrtA_{\Delta N59}$ PM.

In order to assess the thermostability at a temperature higher than 45°C we incubated the mutants for 90 min at 60°C, followed by measurement of the enzymatic activity (Fig. 6). Both single mutants were more resistant to the increased temperature than the WT. Additionally, mutation V193R introduced into the PM made the mutant more thermostable. Unfortunately, mutation M155V did not have a similar effect on thermostability as V193R when introduced into the PM.

Since the catalytic function of $SaSrtA_{\Delta N59}$ has been described to rely on Ca^{2+} ions (Hirakawa *et al.*, 2012; Jeong *et al.*, 2017), we examined the activity of mutants M155V and V193R, both exhibiting higher activity than the WT, in the presence and absence of this divalent ion. As a control for the activity of $SaSrtA_{\Delta N59}$ WT and mutants in the absence of Ca^{2+}, we used a known Ca^{2+}-independent $SaSrtA_{\Delta N59}$ mutant (E105K/E108Q, indicated as DM^{-Ca}) (Hirakawa *et al.*, 2012). As shown in Fig. 7 and in supporting information, Fig. S3, in the absence of the allosteric activator, single mutant

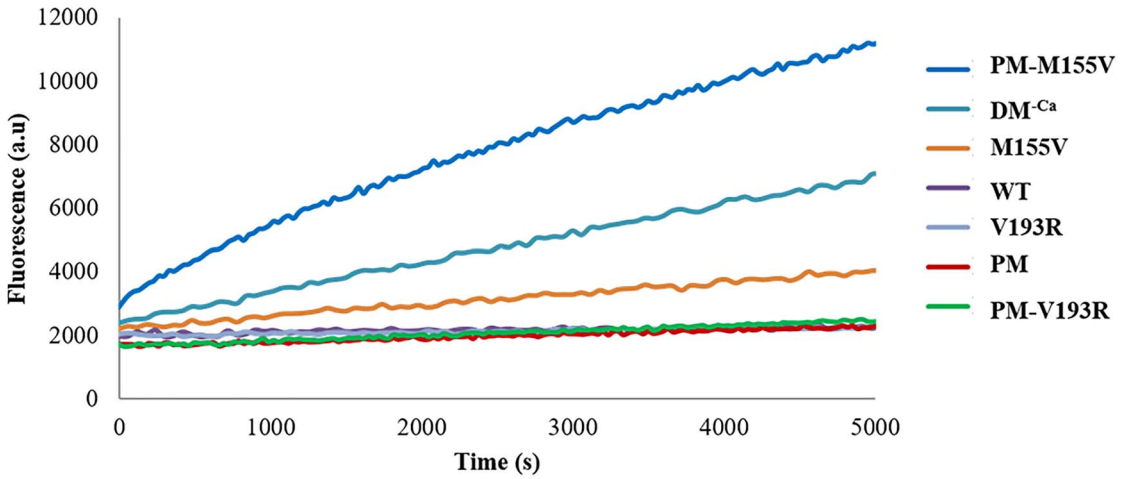

Fig. 7 Representative graphs showing the enzymatic activity of selected variants (1 μM) in the absence of the Ca^{2+} ions (presence of 5 μM EGTA) ($n = 3$).

Table II. Thermodynamic parameters of $SaSrtA_{\Delta N59}$ WT and mutants ($n = 3$, $^*P < 0.01$)

Protein	T_m (°C)	ΔH_u (kJ/mol)	ΔG_u (kJ/mol)
WT	59 ± 0.2	224.9 ± 1.5	4.0 ± 0.2
M155V	56 ± 1.3	259.5 ± 3.5*	8.4 ± 0.4*
V193R	57 ± 0.5	270.8 ± 3.1*	9.7 ± 0.3*
PM	57 ± 0.1	234.5 ± 0.4*	6.9 ± 0.1*
PM-M155V	54 ± 0.9	181.7 ± 3.1	3.1 ± 0.2
PM-V193R	55 ± 0.8	271.6 ± 2.7*	11.1 ± 0.8*

M155V exhibited approximately a 3-fold increase in activity in comparison to the WT. Interestingly, mutation M155V conferred Ca^{2+}-independence to the PM, which may allow the use of this enzyme for *in vivo* conjugation reactions, thus forming an interesting addition to Ca^{2+}-independent sortase A mutants.

Characterization of the thermostability

Mutants M155V, V193R and PM-M155V and PM-V193R exhibiting enzymatic activities higher or comparable to $SaSrtA_{\Delta N59}$ WT were examined for their thermostability using CD. The reversibility of unfolding was monitored by gradually subjecting samples to a temperature of 70°C with steps of 1°C/min, subsequently cooling down to 20°C and repeating the process; as shown previously for WT and PM, the unfolding and refolding transitions were found to be superimposable, which indicates that the proteins exhibited reversible unfolding (Fig. 8).

The fraction folded protein was calculated from the CD measurement and the ellipticity at 210 nm. The reversibility of the temperature-induced unfolding allowed the estimation of thermodynamic parameters from the CD measurement using the Van't Hoff equation. Slopes from the Van't Hoff plots (supplementary information, Fig. S1), which gave R-squared values higher than 0.9, were used for the calculation of the thermodynamic parameters given in Table II.

Although a slight decrease in T_m values of the mutants was observed, all examined variants except PM-M155V showed significantly higher thermodynamic stability with increased ΔH_u and ΔG_u compared to WT (two-tailed Student's t test, $P < 0.01$). The only combination of mutations that led to reduced enzyme stability was found in the PM-M155V mutant. In contrast, the single modification of position M155V led to an increased thermostability of the $SaSrtA_{\Delta N59}$ WT. Nevertheless, the most prominent modification of the $SaSrtA_{\Delta N59}$ WT which increased both ΔH_u and ΔG_u was found at position V193 as a single mutation and in combination with the PM.

Discussion

An overall stabilizing effect of a protein results from the presence of stabilizing and neutral, but also destabilizing mutations, which counterbalance each other (Lehmann *et al.*, 2000). As previously reported, thermophilic proteins have a higher frequency of the residues R and Y in their structures (Kumar *et al.*, 2002) as well as amino acids G, K and I (Farias and Bonato, 2003) or G, A and V (Panja *et al.*, 2015). On the other hand, amino acids N, Q, M and C are considered to bring thermolability and are not often found in thermophilic proteins (Kumar *et al.*, 2002; Taylor and Vaisman, 2010). Nevertheless, thermal stability of a protein is not only dependent on the occurrence of particular amino acids but also on the location of the residue selected for substitution as well as its interactions with other residues (Kumar *et al.*, 2002).

To broaden its range of applications, a protein's stability can be engineered, either via the introduction of appropriate mutations or chemical modification at the surface of the protein (Modarres *et al.*, 2016). Currently, with the advancement of computational techniques and DNA sequencing technologies, sequence-based approaches have gained interest and have led to the introduction of a consensus concept to this field of science (Lehmann *et al.*, 2000; Lehmann and Wyss, 2001; Pantoliano *et al.*, 1989). The utility of this approach has been demonstrated by improving the stabilities of different proteins such as endoglucanases (Anbar *et al.*, 2012), fluorescent proteins (Dai *et al.*, 2007) and a maltogenic amylase (Kim *et al.*, 2003).

Although consensus design has been successfully applied to many proteins to improve their stability (Modarres *et al.*, 2016), some key factors need to be considered before using this methodology. One of the most important aspects is the generation of an MSA. It requires the utilization of high-quality sequences with well-conserved regions, similar length and no repetitions, which are not always available for the protein of interest (Magliery *et al.*, 2011).

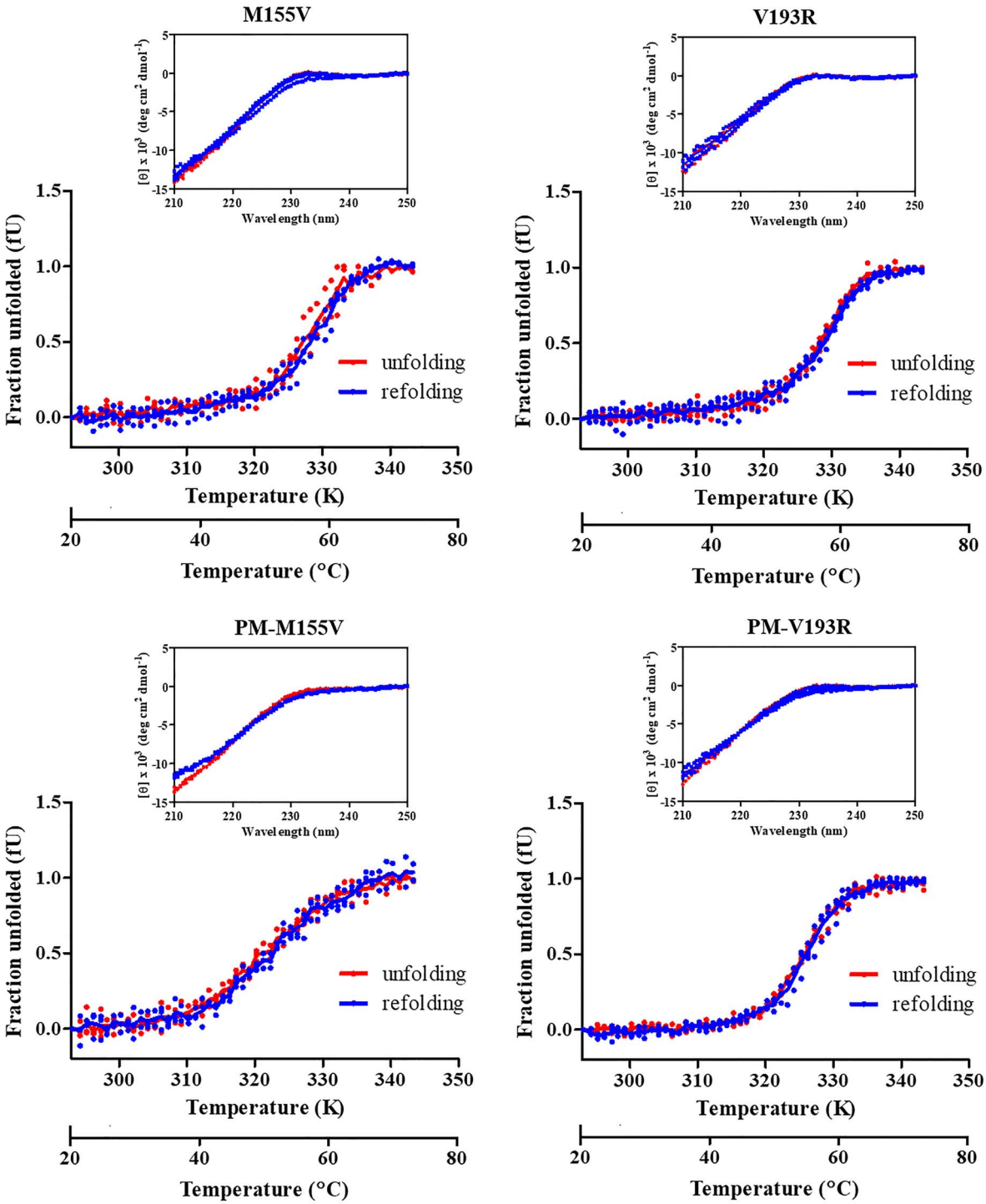

Fig. 8 Reversibility of protein unfolding as measured by CD ($n = 3$). Individual graphs represent the unfolding of native (red line) and unfolding of refolded (blue line) single mutants M155V and V193R and PM-M155V and PM-V193R (Romero-Romero *et al.*, 2015). The inset figures show the CD spectra of the native (red) and refolded proteins (blue).

Additionally, the application of sequences belonging to different taxonomic groups can decrease the bias towards specific taxa or species. Apart from the quality, the number of sequences can also influence the accuracy of the MSA, although an optimal number is not established (Jacobs *et al.*, 2012; Porebski *et al.*, 2015). In other words, both the quality and the quantity of sequences used for the generation of the MSA need to be properly selected. Our approach to the improvement of SaSrtA's thermostability covers more than a hundred homologous input sequences belonging to nine families of bacteria available. Before generating the MSA, all duplicates were removed, and sequences were evaluated for their length.

In order to get a complete picture of SaSrtA's stability, we calculated the T_m value, which is most commonly used for the estimation of thermal resistance. In addition, we determined ΔG_u values, which take the difference in free energy of the folded and unfolded species into consideration (Sanfelice and Temussi, 2016; Wright *et al.*, 2017). Here, we corroborate that the structure of $SaSrtA_{\Delta N59}$ WT is thermodynamically stable and exhibits a reversible unfolding (Fig. 1). This is a great evolutionary advantage, which helps prevent protein aggregation in unfavorable conditions (Sanfelice and Temussi, 2016). We decided to engineer the $SaSrtA_{\Delta N59}$ WT and the $SaSrtA_{\Delta N59}$ PM to potentially improve their thermostability: detection of destabilizing mutations and their removal can lead to an increase in thermostability (Lehmann *et al.*, 2000). We identified two thermodynamically more stable mutations, M155V and V193R (Fig. 4 and Table II). The V193R mutation had in fact the biggest impact on the thermostability of SaSrtA and was found via strategy 2, for which only extremophiles were used to build the MSA. In comparison to WT, both M155V and V193R showed a more than 2-fold improved ΔG_u, which could reduce the loss of activity over time. Indeed, mutants M155V and V193R did show a higher residual activity after incubation at 60°C compared to WT and even PM (Fig. 6). One would expect a higher enzymatic activity at higher temperature (Arcus *et al.*, 2016); interestingly, the mutants (alone and in combination with the PM) did not show an increased initial reaction rate (Fig. 5). However, not all enzymes exhibit perfect Michaelis-Menten kinetics (Daniel and Danson, 2013). Although ΔG_u improved, ΔH_u and T_m did not undergo a similar improvement (Table II); we speculate this combination could lead to an improved stability without a further improvement in the enzymatic activity (Daniel and Danson, 2013).

Ca^{2+} is known to enhance the activity of SaSrtA (Ilangovan *et al.*, 2001; Naik *et al.*, 2006; Suree *et al.*, 2009). Binding of the ion is coordinated through E105, E108 and E171, residues that are remote from the active site. Ca^{2+} binding combined with substrate binding leads to a closed conformation, which involves a disorder-to-order transition of the $\beta6/\beta7$ loop (Suree *et al.*, 2009; Wang *et al.*, 2018) and extensive structural changes such as shortening of part of the loop (T156 to V161). Previously, Ca^{2+}-independence was achieved by introducing mutations E105K and E108Q (Hirakawa *et al.*, 2012). These mutations were also introduced into the PM; though the mutants were active in the absence of Ca^{2+}, the activity was generally less than in the presence of the allosteric activator (Jeong *et al.*, 2017; Witte *et al.*, 2015; Wuethrich *et al.*, 2014). Interestingly, we found that mutation M155V rendered the enzyme active in the absence of Ca^{2+} (Fig. 7), even though M155V is located at the C-terminal end of strand $\beta6$ lining the substrate groove and not involved with Ca^{2+} binding. Ca^{2+}-independence was conferred to PM-M155V as well. We hypothesize that mutations lining the substrate groove, such as M155V but also D160A and D165A in the PM, may improve substrate binding by altering the conformation of the $\beta6/\beta7$ loop, thereby also changing the role for the allosteric activator.

In summary, we explored the thermodynamic stability of *S. aureus* SrtA and potential improvement thereof by means of consensus design. Consensus design was successfully applied to identify specific residues in the $SaSrtA_{\Delta N59}$ WT sequence that did not only have an impact on thermal stability but also on enzymatic activity and even on Ca^{2+}-dependence. We therefore anticipate that variants such as M155V and PM-M155V could find an application in conjugation reactions in low Ca^{2+} environments and potentially be used in *in vivo* labeling experiments.

Supplementary data

Supplementary data are available at *PEDS* online.

Acknowledgments

We would like to thank Dr T.E. Adams (CSIRO Manufacturing, Parkville, Australia) for kindly providing the *srtA* gene. We acknowledge the support of Dr S. Romero-Romero with the thermodynamics experiments. The authors would also like to thank Dr C. Mayer for his help with CD experiments. The paper is edited by Dr Alan Berry, PEDS Board Member.

Funding

Human Frontiers Science Program (HFSP) long-term fellowship (LT001131/2011) and a Rosalind Franklin Fellowship (University of Groningen) to Y.L.B.; Orange Tulip Scholarship provided to S.V.T.

References

An, C., Chaikof, E.L., Ham, H.O., Liu, D.R., Dorr., B.M. (2014) *Proc. Natl. Acad. Sci.*, **111**, 13343–13348.
Anbar, M., Gul, O., Lamed, R., Sezerman, U.O., Bayer, E.A. (2012) *Appl. Environ. Microbiol.*, **78**, 3458–3464.
Antos, J.M., Truttmann, M.C., Ploegh, H.L. (2016) *Curr. Opin. Struct. Biol.*, **38**, 111–118.
Arcus, V.L., Prentice, E.J., Hobbs, J.K., Mulholland, A.J., Van der Kamp, M.W., Pudney, C.R., Parker, E.J., Schipper, L.A. (2016) *Biochemistry*, **55**, 1681–1688.
Arnold, F.H. and Volkov, A.A. (1999) *Curr. Opin. Chem. Biol.*, **3**, 54–59.
Bentley, M.L., Lamb, E.C., McCafferty, D.G. (2008) *J. Biol. Chem.*, **283**, 14762–14771.
Call, E.K. and Klaenhammer, T.R. (2013) *Front. Microbiol.*, **4**, 1–10.
Chen, R. (2001) *Trends Biotechnol.*, **19**, 13–14.
Chen, I., Dorr, B.M., Liu, D.R. (2011) *Proc. Natl. Acad. Sci.*, **108**, 11399–11404.
Chen, L., Cohen, J., Song, X., Zhao, A., Ye, Z., Feulner, C.J., Doonan, P., Somers, W., Lin, L., Chen, P.R. (2016) *Sci. Rep.*, **6**, 1–12.
Cobb, R.E., Chao, R., Zhao, H. (2013) *AICHE J.*, **59**, 1432–1440.
Dai, M., Fisher, H.E., Temirov, J., Kiss, C., Phipps, M.E., Pavlik, P., Werner, J.H., Bradbury Protein Eng, A.R. (2007) *Des Sel*, **20**, 69–79.
Daniel, R.M. and Danson, M.J. (2013) *FEBS Lett.*, 2738–2743.
Farias, S.T. and Bonato, M.C.M. (2003) *Genet. Mol. Res.*, **2**, 383–393.
Frankel, B.A., Tong, Y., Bentley, M.L., Fitzgerald, M.C., McCafferty, D.G. (2007) *Biochemistry*, 7269–7278.
Greenfield, N.J. (2006) *Nat. Protoc.*, **1**, 2876–2890.
Hirakawa, H., Ishikawa, S., Nagamune, T. (2012) *Biotechnol. Bioeng.*, **109**, 2955–2961.
Ilangovan, U., Ton-That, H., Iwahara, J., Schneewind, O., Clubb, R.T. (2001) *Proc. Natl. Acad. Sci.*, **98**, 6056–6061.
Jäckel, C., Bloom, J.D., Kast, P., Arnold, F.H., Hilvert, D. (2010) *J. Mol. Biol.*, **399**, 747–762.
Jacobitz, A.W., Kattke, M.D., Wereszczynski, J., Clubb, R.T. (2017) *Adv. Protein Chem. Struct. Biol.*, **155**, 223–264.
Jacobs, S.A., Diem, M.D., Luo, J., Teplyakov, A., Obmolova, G., Malia, T., Gilliland, G.L., O'Neil, K.T. (2012) *Protein Eng. Des. Sel.*, **25**, 107–117.
Jeong, H.J., Abhiraman, G.C., Story, C.M., Ingram, J.R., Dougan, S.K. (2017) *PLoS One*, **12**, 1–15.
Kim, Y.W., Choi, J.H., Kim, J.W., Park, C., Kim, J.W., Cha, H., Lee, S.B., Oh, B.H., Moon, T.W., Park, K.H. (2003) *Appl. Environ. Microbiol.*, **69**, 4866–4874.
Klotz, I.M. (1996) *Proc. Natl. Acad. Sci.*, **93**, 14411–14415.
Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L.L., Mol, E.L.L.J. (2001) *Biol.*, **305**, 567–580.
Kumar, S., Tsai, C.J., Nussinov, R. (2002) *Protein Eng. Des. Sel.*, **13**, 179–191.

Lehmann, M. and Wyss, M. (2001) *Curr. Opin. Biotechnol.*, **12**, 371–375.
Lehmann, M., Pasamontes, L., Lassen, S.F., Wyss, M. (2000) *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.*, **1543**, 408–415.
Li, Z. and Hirst, J.D. (2017) *Chem. Sci.*, **8**, 4318–4333.
Liu, H. and Naismith, J.H. (2008) *BMC Biotechnol.*, **8**, 91.
Liu, F., Du, D., Fuller, A.A., Davoren, J.E., Wipf, P., Kelly, J.W., Gruebele, M. (2008) *Proc. Natl. Acad. Sci.*, **105**, 2369–2374.
Lutz, S. (2011) *Curr. Opin. Biotechnol.*, **21**, 734–743.
Magliery, T.J., Lavinder, J.J., Sullivan, B.J. (2011) *Curr. Opin. Chem. Biol.*, **15**, 443–451.
Marraffini, L.A., DeDent, A.C., Schneewind, O. (2006) *Microbiol. Mol. Biol. Rev.*, **70**, 192–221.
Mazmanian, S.K., Liu, G., Ton-That, H., Schneewind, O. (1999) *Science*, **285**, 760–763.
Milardi, D., La Rosa, C., Fasone, S., Grasso, D. (1997) *Biophys. Chem.*, **69**, 43–51.
Modarres, H.P., Mofrad, M.R., Sanati-Nezhad, A. (2016) *RSC Adv.*, **6**, 115252–115270.
Naik, M.T., Suree, N., Ilangovan, U., Liew, C.K., Thieu, W., Campbell, D.O., Clemens, J.J., Jung, M.E., Clubb, R.T. (2006) *J. Biol. Chem.*, 1817–1826.
Panja, A.S., Bandopadhyay, B., Maiti, S. (2015) *PLoS One*, **10**, 1–21.
Pantoliano, M.W., Whitlow, M., Wood, J.F., Dodd, S.W., Hardman, K.D., Rollence, M.L., Bryan, P.N. (1989) *Biochemistry*, **28**, 7205–7213.
Pei, J. and Grishin, N.V. (2001) *Bioinformatics*, **17**, 700–712.
Pei, J., Kim, B.H., Grishin, N.V. (2008a) *Nucleic Acids Res.*, **36**, 2295–2300.
Pei, J., Tang, M., Grishin, N.V. (2008b) *Nucleic Acids Res.*, **36**, 30–34.
Pelay-Gimeno, M., Bange, T., Hennig, S., Grossmann, T.N. (2018) *Angew. Chem.*, **57**, 11164–11170.
Porebski, B.T. and Buckle, A.M. (2016) *Protein Eng. Des. Sel.*, **29**, 245–251.
Porebski, B.T., Nickson, A.A., Hoke, D.E., Hunter, M.R., Zhu, L., McGowan, S., Webb, G.I., Buckle, A.M. (2015) *Protein Eng. Des. Sel.*, **28**, 67–78.
Proft, T. (2009) *Biotechnol. Lett.*, **32**, 1–10.
Rampelotto, P.H. (2013) *Lifestyles*, **3**, 482–485.
Romero-Romero, S., Costas, M., Rodríguez-Romero, A., Fernández-Velasco, D.A. (2015) *Phys. Chem. Chem. Phys.*, **17**, 20699–20714.
Sanfelice, D. and Temussi, P.A. (2016) *Biophys. Chem.*, **208**, 4–8.
Steiner, K. (2012) *Biotechnol. J.*, **2**, e201209010.
Steiner, K. and Schwab, H. (2012) *Comput. Struct. Biotechnol. J.*, **2**, e201209010.
Steipe, B., Schiller, B., Pluckthun, A., Steinbacher, S. (1994) *J. Mol. Biol.*, **240**, 188–192.
Suree, N., Liew, C.K., Villareal, V.A., Thieu, W., Fadeev, E.A., Clemens, J.J., Jung, M.E., Clubb, R.T. (2009) *J. Biol. Chem.*, **284**, 24465–24477.
Taylor, T.J. and Vaisman, I.I. (2010) *BMC Struct. Biol.*, **10**, S5.
Ton-That, H., Mazmanian, S.K., Faull, K.F., Schneewind, O. (2000) *J. Biol. Chem.*, **275**, 9876–9881.
Tripathi, T. (2013) *J Proteins and Proteomics*, **4**, 85–91.
UniProt Consortium (2009) *Nucleic Acids Res.*, **37**, D169–D174.
Wang, X., Chen, J.L., Otting, G., Su, X.C. (2018) *Sci. Rep.*, 2–10.
Witte, M.D., Wu, T., Guimaraes, C.P., Theile, C.S., Blom, A.E., Ingram, J.R., Li, Z., Kundrat, L., Goldberg, S.D., Ploegh, H.L. (2015) *Nat. Protoc.*, 508–516.
Wójcik, M., Eleftheriadis, N., Zwinderman, M.R.H., Dömling, A.S.S., Dekker, F.J., Boersma, Y.L. (2019) *Eur. J. Med. Chem.*, **161**, 93–100.
Wong, H.J., Stathopulos, P.B., Bonner, J.M., Sawyer, M., Meiering, E.M. (2004) *J. Mol. Biol.*, **344**, 1089–1107.
Wright, T.A., Stewart, J.M., Page, R.C., Konkolewicz, D. (2017) *J. Phys. Chem. Lett.*, **8**, 553–558.
Wu, Q., Ploegh, H.L., Truttmann, M.C. (2017a) *ACS Chem. Biol.*, **12**, 664–673.
Wu, Z., Hong, H., Zhao, X., Wang, X. (2017b) *Bioresour Bioprocess*, 4.
Wuethrich, I., Peeters, J.G.C., Blom, A.E.M., Theile, C.S., Li, Z., Spooner, E., Ploegh, H.L., Guimaraes, C.P. (2014) *PLoS One*, **9**.
Xu, B.L., L, B., Dai, M., Chen, Y., Meng, D., Fang, Y.W.N., Tang, X.F., Tang, B. (2015) *Appl. Environ. Microbiol.*, **81**, 6302–6313.