

Supplementary Information

Directed sortase A evolution for efficient site-specific bioconjugations in organic co-solvents

Zhi Zou,^{ab} Hoda Alibiglou,^b Diana M. Mate,^a Mehdi D. Davari,^b Felix Jakob,^{ab}
and Ulrich Schwaneberg^{*ab}

^aDWI - Leibniz-Institute for Interactive Materials, Forckenbeckstraße 50, 52056 Aachen, Germany.

^bInstitute of Biotechnology, RWTH Aachen University, Worringerweg 3, 52074 Aachen, Germany.

*Correspondence to: Prof. Dr. Ulrich Schwaneberg, RWTH Aachen University, Institute of Biotechnology, Worringerweg 3, 52074 Aachen. Tel.: +49 241 80 24170. Fax: +49 241 80 22578. E-mail: u.schwaneberg@biotec.rwth-aachen.de

Table of Contents

Directed sortase A evolution for efficient site-specific bioconjugations in organic co-solvents	1
Section 1: Engineering of Sa-SrtA towards organic co-solvents.....	4
Experimental part.....	4
Materials.....	4
Optimization of SortEvolve screening assay in DMSO co-solvent.....	4
Determination of coefficient of variation of SortEvolve assay in 45% (v/v) DMSO co-solvent	5
KnowVolution of Sa-SrtA towards organic solvents.....	5
Characterization of improved Sa-SrtA variants in absence and presence of 45% (v/v) DMSO.....	7
Activity profiles of Sa-SrtA in different organic co-solvents.....	7
Sortase-mediated protein-protein bioconjugation in DMSO and ethanol co-solvents	8
List of supplementary Tables.....	9
Table. S1	9
Table. S2	9
Table. S3	10
Table. S4	11
Table. S5	12
Table. S6	12
List of supplementary Figures	13
Fig. S1.....	13
Fig. S2.....	14
Fig. S3.....	15
Fig. S4.....	16
Fig. S5.....	17
Fig. S6.....	18
Fig. S7.....	19
Section 2: Computational study of Sa-SrtA WT and variants in DMSO co-solvent	20
<i>In silico</i> generation of sortase variants.....	20
Molecular dynamics (MD) simulations.....	20
List of supplementary Figures	22
Fig. S8.....	25
Fig. S9.....	26
Fig. S10.....	27
Fig. S11.....	28

Fig. S12.....	29
Fig. S13.....	30
Fig. S14.....	31
Fig. S15.....	32
Fig. S16.....	33
Fig. S17.....	34
Fig. S18.....	35
Fig. S19.....	36
Section 3: Application of engineered sortase A variants for peptide semi-synthesis in organic co-solvents.....	37
Materials.....	37
Experimental part.....	37
Sortase-mediated ligation of hydrophobic antiviral peptide AVP 0683 and Abz-LPETGK-Dnp in DMSO or DMF co-solvent.....	37
Sortase-mediated ligation of hydrophobic amines (tyramine or 4-(Trifluoromethyl)-benzylamine (4-TFB amine)) and Abz-LPETGK-Dnp in DMSO co-solvent.....	38
Sortase-mediated PEGylation of hydrophobic peptide Abz-LPETGK-Dnp in DMSO co-solvent....	38
List of supplementary Table	39
Table. S7	39
List of supplementary Figures	40
Fig. S20.....	40
Fig. S21.....	41
Fig. S22.....	42
Fig. S23.....	43
Fig. S24.....	44
Fig. S25.....	45

Section 1: Engineering of Sa-SrtA towards organic co-solvents

Experimental part

Materials

Chemical reagents and solvents with analytical grade or higher purity were purchased from Sigma-Aldrich (Hamburg, Germany), AppliChem (Darmstadt, Germany), Carl Roth (Karlsruhe, Germany). Dimethylsulfoxide (DMSO, 99.5%, AppliChem), dimethylformamide (DMF, 99.5%, AppliChem), methanol (99.8%, Sigma-Aldrich), ethanol (99.8%, Sigma-Aldrich) were purchased. Peptides (or peptide derivatives) Abz-LPETGK-Dnp-NH₂ (97.8%), Abz-LPETGGG-COOH (97.2%), antiviral peptide 1 and antiviral peptide 2 were purchased from Bachem (Bubendorf, Switzerland). Enzymes were all purchased from New England Biolabs (Frankfurt, Germany) or Fermentas (St. Leon-Rot, Germany). Primers used in polymerase chain reactions (PCR) were purchased from Eurofins MWG Operon. V/flat-bottom polystyrene 96-well microtiter plates and flat-bottom polypropylene 96-well microtiter plates were purchased from Greiner Bio-One GmbH, (Frickenhausen, Germany).

Optimization of SortEvolve screening assay in DMSO co-solvent

Determination of GGG-eGFP-LCI resistance in DMSO co-solvents

GGG-eGFP-LCI was expressed and purified as reported.¹ Resistance of GGG-eGFP-LCI in DMSO co-solvents was evaluated by the fluorescence intensity of eGFP in presence of different DMSO concentrations. In detail, purified GGG-eGFP-LCI was incubated (3 h, 600 rpm, room temperature, MTP shaker, TiMix5, Edmund Bühler GmbH, Hechingen, Germany) with gradient concentration of DMSO (0 to 80% (v/v)) in buffer A (100 μ L, 5 mM CaCl₂, 150 mM NaCl, 50 mM Tris-HCl, pH 7.5) in polypropylene microtiter plates (PP-MTP). Subsequently eGFP fluorescence was determined (Tecan infinite 1000Pro plate reader, λ_{exc} = 488 nm; λ_{em} = 509 nm, gain = 100; Tecan Group AG, Männedorf, Switzerland). The determined fluorescence of GGG-eGFP-LCI is shown in **Fig. S1a**.

Determination of CueO-LPETGGRR resistance in DMSO co-solvent

Purified CueO-LPETGGRR was produced as reported.¹ Resistance of CueO-LPETGGRR against DMSO co-solvents was evaluated by a standard 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid (ABTS, ϵ = 36000 M⁻¹ cm⁻¹) assay. In detail, purified CueO-LPETGGRR (50 μ g/mL) was incubated (3 h, 600 rpm, room temperature, MTP shaker, TiMix5) with gradient concentration of DMSO (0 to 80% (v/v)) in 100 μ L buffer A in polypropylene microtiter plates (PP-MTP). Ten microliter liquid was transferred into 190 μ L buffer B (100 mM, pH 3.0, sodium citrate) with 3 mM ABTS. Plates were stirred for 5 seconds and the absorbance was determined (420 nm (ϵ_{ABTS^+} , 420 nm = 36,000 M⁻¹cm⁻¹), room temperature, Tecan Infinite 1000Pro plate reader). The determined activity of CueO-LPETGGRR is shown in **Fig. S1b**.

Optimization of DMSO concentration in the SortEvolve screening assay

A sortase screening assay “SortEvolve” was performed in different concentrations of DMSO with a modified protocol based on Application.¹ Five main steps are included. In *Step1*, Sa-SrtA WT was expressed in MTP and clear supernatant of cell lysates were obtained as previously described.¹ Sa-SrtA containing cell-free lysates (30 μ L per well) were transferred into F-bottom 96-well PS-MTPs. In *Step2*, conjugation of CueO-LPETGGGRR and GGG-eGFP-LCI was performed in F-bottom 96-well PS-MTPs (reaction mixture: 200 μ L, 100 μ g/mL purified GGG-eGFP-LCI, 50 μ g/mL purified CueO-LPETGGGRR, different concentrations of DMSO (range from 0 to 55% (v/v) in buffer A) followed by an incubation (800 rpm, 3 h, room temperature). *Step3*, *4* and *5* were performed as reported.¹ Residual activity of SortEvolve assay in different concentrations of DMSO is presented in **Fig.S1c**). Based on the curve, a linear range of decreased activity of SortEvolve assay was observed when 35 to 55% (v/v) DMSO was incubated. Consequently, 45% (v/v) DMSO was selected for screening of Sa-SrtA variants.

Determination of coefficient of variation of SortEvolve assay in 45% (v/v)

DMSO co-solvent

In order to determine the coefficient of variation of SortEvolve assay in buffer and in presence of 45% (v/v) DMSO, two 96-well MTPs containing Sa-SrtA WT were screened (in order to gain information on background, six wells contained an “empty vector” control and six wells contained the TB-expression media). Slopes of ABTS absorbance over time were analyzed. The overall activity of SortEvolve in 45% (v/v) DMSO was 60% decreased compared to activity in buffer. Coefficient of variation of SortEvolve assay in absence and presence of 45% (v/v) DMSO were calculated as 12.9% and 14.5%, respectively (**Fig. S1d**).

KnowVolution of Sa-SrtA towards organic solvents

The directed Sa-SrtA evolution for organic solvents was performed with a standard *KnowVolution* strategy.²

Diversity generation of Sa-SrtA library

A sequence saturation mutagenesis (SeSaM) library of Sa-SrtA lacking the N-terminal 59 residues (PDB code: 2KID) was generated.³ Unless otherwise stated, the standard PCRs and the related primers in PCRs during the generation of SeSaM library were performed as reported.⁴ Templates for individual steps were generated (**Fig. S2a**).⁵ The following phosphorothioate deoxynucleotides (dATP α S and dGTP α S) concentrations were used: A-forward library-35% (**Fig. S2b**), A-reverse library-35% (**Fig. S2c**), G-forward library-40% (**Fig. S2d**), and G-reverse library-40% (**Fig. S2e**). The final Sa-SrtA-SeSaM library was generated using 200 ng PCR products of each library (**Fig. S2f**) and cloned into pET28a(+) vector via phosphorothioate-based ligase-independent gene cloning (PLICing).⁶ The PCR conditions for amplification of vector backbone for PLICing were 98°C for 60 sec, 98°C for 45 sec, 55°C for 30 sec, 72°C for 3 min (25 cycles); 72°C for 10 min (one cycle). The PCR conditions for amplification of insert (Sa-SrtA-SeSaM library) for PLICing were 98°C for 60 sec, 98°C for 45 sec, 55°C for 30 sec, 72°C for 45 sec (25 cycles); 72°C for 5 min (1 cycle). Primers for PCRs are listed in **Table. S1**. The PLICing product of Sa-SrtA-SeSaM was transformed into chemically competent *E. coli* BL21 Gold (DE3). Colonies (on agar plate after transformation) from the generated SeSaM library were randomly selected for sequencing. Four Sa-SrtA wide-types were found among 13 selected colonies. Eleven mutations were found in 9 variants. Transition and transversion rates of mutations were calculated as 55% and 45%, respectively.

Phase I: Screening of SeSaM library in DMSO co-solvent

The generated library (1680 colonies) was transferred in 96-well polystyrene microtiter plates (V-bottom, one clone per well). Each plate contained six wells with negative control (cells containing empty vector pET-28 instead of Sa-SrtA). Protein was expressed and lysate was produced as previously described.¹ For the library screening in DMSO co-solvent, each plate was screened with the optimized protocol. In brief, a five-step work flow was followed. *Step1*, an aliquot (30 μ L) of library lysate from each well was transferred into in 96-well polystyrene microtiter plates (F-bottom). *Step2* Conjugation of CueO-LPETGGRR and GGG-eGFP-LCI was performed in 96-well polystyrene microtiter plates (F-bottom, reaction mixture: 200 μ L, 100 μ g/mL purified GGG-eGFP-LCI, 50 μ g/mL purified CueO-LPETGGRR, 45% (v/v) of DMSO, in buffer A) followed by incubation (800 rpm, 3 h, room temperature). *Step3*, *4* and *5* were performed as previously described.¹ The activity values from negative controls were averaged as the background and subtracted in all cases. Variants with 1.16-fold or higher improved activity (compared to Sa-SrtA WT) were selected for subsequent re-screening. Rescreening was performed with six replicates per clone using the aforescribed screening protocol. Rescreening results of the SeSaM library variants are showed **Fig. S3a**.

Phase II: generation and screening of site-saturation mutagenesis

SSM libraries at positions P94 and D165 were generated as previously described.¹ Primers used in new generated SSM libraries are shown in **Table. S2**. *Fw SSM* and *Rev SSM* primers were used to generate Sa-SrtA at the corresponding positions. *Fw SSM D186/K196* and *Rev SSM D186/K196* were used to saturate simultaneously two sites in Sa-SrtA _{Δ 59}. To generate all SSM PCRs, the following protocol was used: 98°C for 45 sec (1 cycle); 98°C for 45 sec, 58°C for 30 sec, 72°C for 3 min 30 sec (25 cycles); 72°C for 10 min (1 cycle). PCR solutions (50 μ L) for amplification consist of plasmid template (15 ng), dNTP mix (10 mM), PfuS DNA polymerase (2.5 U), and forward and reverse primer (50 μ M each). Parental DNA was digested by *Dpn* I (5 U, 37°C, overnight). *Dpn*I was heat inactivated (80°C for 20 min). PCR products were purified (PCR clean-up kit, Macherey-Nagel™, Düren, Germany) subsequently transformed into *E.coli* BL-21(DE3) competent cells. One hundred and sixty-eight clones of each single-site SSM library were transferred into two 96-well polystyrene microtiter plates (V-bottom) and 504 clones of double-site SSM-D186/K196 were transferred in to six 96-well polystyrene microtiter plates (V-bottom). Cells were cultivated and lysate containing Sa-SrtA was produced using the methods described above. The aforescribed protocol for screening of SeSaM library was employed for the screening and rescreening of all the SSM libraries. Rescreening results of the identified variants from all single-site SSM libraries are summarized in **Fig. S3c**. Sequencing results of the identified variants are summarized in **Table S3**. Rescreening results of variants identified in the SSM-D186/K196 library are summarized in **Fig. S3d**.

Phase IV: recombination

Recombination of identified amino acid substitutions was conducted via site-directed mutagenesis (SDM). Primers used in SDM are listed in **Table.S4**. In detail, the isolated variant Sa-SrtA D186G/K196V (M2) was used as the template to generated variants Sa-SrtA R159G/D186G/K196V (primers *Fw SDM R159G* and *Rev SDM R159G* were employed), Sa-SrtA R159T/D186G/K196V (primers *Fw SDM R159T* and *Rev SDM R159T* were employed), Sa-SrtA D165A/D186G/K196V (primers *Fw SDM D165A* and *Rev SDM D165A* were employed), Sa-SrtA D165Q/D186G/K196V (primers *Fw SDM D165Q* and *Rev SDM D165Q* were employed), Sa-SrtA D170W/D186G/K196V (primers *Fw SDM D170W* and *Rev SDM D170W* were employed) and Sa-SrtA

P94S/D160N/D165A/K196T (rM4)⁷ (primers *Fw SDM D160N/D165A*, *Rev SDM D160N/D165A*, *Fw SDM K196T* and *Rev SDM K196T* were employed first step) were employed). The PCR and subsequent PCR product purification were performed as described above. PCR products were purified (PCR clean-up kit, Macherey-Nagel) and subsequently transformed into *E.coli* BL-21(DE3) competent cells.

Fluorimetric assay of recombined Sa-SrtA variants

Recombined variants were expressed in flask and cell-free lysate containing Sa-SrtAs were produced as previously described.¹ A fluorimetric assay was employed to determine Sa-SrtA activity.⁸ In short, reactions (reaction mix: 100 μ L, 0.05 mM Abz-LPETGK-Dnp, 5 mM glycine-glycine-glycine (tri-glycine) in buffer A (buffer A: 5 mM CaCl₂, 150 mM NaCl, 50 mM Tris-HCl, pH 7.5)) were initiated by adding 5 μ L Sa-SrtAs cell-free lysate. The increase in fluorescence was continuously detected (λ_{exc} = 320 nm; λ_{em} = 420 nm, gain = 100, Tecan infinite 1000Pro plate reader). The activities of selected variants are shown in **Table. S5**.

Characterization of improved Sa-SrtA variants in absence and presence of 45% (v/v) DMSO

Kinetics of Sa-SrtAs was evaluated via a HPLC assay which was previously described.⁷ Reactions (50 μ L) were performed with 9 mM NH₂-Gly-Gly-Gly-COOH, varied concentrations of Abz-LPETGK-Dnp-NH₂ (0.25 to 5 mM) and 2 μ M Sa-SrtAs in buffer A (5 mM CaCl₂, 150 mM NaCl, 50 mM Tris-HCl buffer, pH 7.5), or 45% (v/v) DMSO co-solvent (22.5 μ L DMSO, 5 mM CaCl₂, 150 mM NaCl, 50 mM Tris-HCl buffer pH 7.5). Reactions were performed in room temperature for 5 to 60 min before quenching with 25 μ L HCl (1 M). The quenched reaction mixture was then diluted 5-times (final volume: 375 μ L) with pure water. Twenty microliter of the diluted sample was injected into a reversed-phase C18 HPLC column (4.6x150 mm, 5 μ M, Macherey-Nagel, Düren, Germany) and chromatographed using a gradient of 10 to 40% acetonitrile with 0.1% TFA (trifluoroacetic acid) in 0.1% aqueous TFA over 20 minutes. Retention times for Abz-LPETGK-Dnp-NH₂, GK-Dnp-NH₂ and Abz-LPETGGG-COOH were 15.2, 13.2 and 11.1 min, respectively. Dnp containing peaks were detected at 355 nm and Abz containing peaks were detected at 255 nm. The yield of product Abz-LPETGGG-COOH was calculated by integrating the area under HPLC trace. K_m and k_{cat} were calculated using Originpro 8.6 (OriginLab, Northampton, USA). **Fig. S4** and **S5** show the plots for the calculation of $K_{m(LPETG)}$ and k_{cat} of variant in absence (buffer A) and in presence of 45% (v/v) DMSO. All the data are summarized in **Table. 1** in the main article.

Activity profiles of Sa-SrtA in different organic co-solvents

A fluorimetric assay⁸ was employed to detect Sa-SrtAs activity in DMSO, DMF, ethanol and methanol co-solvents. In short, reactions (100 μ L, 0.1 mM Abz-LPETGK-Dnp, 5 mM tri-glycine in buffer A (buffer A: 5 mM CaCl₂, 150 mM NaCl, 50 mM Tris-HCl, pH 7.5)) with gradient concentration of DMSO, DMF, ethanol or methanol co-solvents were initiated by supplementing 1.6 μ M purified Sa-SrtA enzyme (WT or variants). The fluorescence of the reaction liquid was constantly determined (λ_{exc} = 320 nm; λ_{em} = 420 nm, gain = 100; Tecan infinite 1000Pro plate reader). The residual activity of Sa-SrtAs in co-solvents was calculated as the ratio of activity in presence of solvent divided by activity in absence of solvent. The relative activity in co-solvent was calculated as the ratio of Sa-SrtA variant's activity divided by Sa-SrtA WT's activity. Residual activities of Sa-

SrtAs in gradient concentration of co-solvents are presented in **Fig. S6**. The relative activities of Sa-SrtAs in 45% (v/v) DMSO, 30% (v/v) DMF, 30% (v/v) ethanol and 50% (v/v) methanol are summarized in **Table. S6**.

Sortase-mediated protein-protein bioconjugation in DMSO and ethanol co-solvents

The protein-protein ligation of CueO-LPETGGGRR and GGG-eGFP-LCI catalyzed by Sa-SrtA was performed in 45% (v/v) DMSO and 30% (v/v) ethanol. In short, the reaction mixture (500 μ L, purified GGG-eGFP-LCI (500 μ g/mL), CueO-LPETGGGRR (500 μ g/mL), purified Sa-SrtA (30 μ g/mL), 45% (v/v) DMSO) was incubated 14 h (room temperature, 800rpm). After ligation, an aliquot (15 μ L) of reaction mixture was first diluted with 60 μ L pure water and subsequently mixed with 25 μ L 4x SDS loading buffer. The mixture was immediately incubated at 95°C for 5 min. Ten microliters of the samples were loaded and analyzed on 10% acrylamide gels. Similarly, reaction mixture (500 μ L, purified GGG-eGFP-LCI 500 μ g/mL, 250 μ g/mL purified CueO-LPETGGGRR, 30 μ g/mL purified Sa-SrtA, 30% (v/v) ethanol) was incubated (room temperature, 800 rpm, 14 h). An aliquot (15 μ L) of reaction mixture was first diluted to 75 μ L with pure water and immediately mixed with 25 μ L 4x SDS loading buffer. The mixture was heated at 95°C for 5 min. Ten microliters of the samples were loaded and analyzed on 10% acrylamide gels. SDS-PAGE of sortase mediated bioconjugation of CueO-LPETGGGRR and GGG-eGFP-LCI is shown in **Fig. S7**.

List of supplementary Tables

Table. S1 List of primers for Sa-SrtA-SeSaM library generation.

Primer name	Sequence (5'-3')
F1	CGACTCACTATAGGGGAATTGTGAGCGGA
R3	CGGGCTTTGTTAGCAGCCGGATCTCAG
SeSaM_F	CACACTACCGCACTCCGTCG
SeSaM_R	GTGTGATGGCGTGAGGCAGC
SeSaM_F1	CACACTACCGCACTCCGTCGCGACTCACTATAGGGGAATTGTGAGCGGA
SeSaM_R3	GTGTGATGGCGTGAGGCAGCCGGCTTTGTTAGCAGCCGGATCTCAG
F1_up	CGCCTGTCACCGACTCACTATAGGGGAATTGTGAGCGGA
R3_dn	GCGGACAGTGCGGGCTTTGTTAGCAGCCGGATCTCAG
Bio_SeSaM_F	[Biotin]CACACTACCGCACTCCGTCG
Bio_SeSaM_R	[Biotin]GTGTGATGGCGTGAGGCAGC
V-F_PLIC	catccgagttcGAAAAGTAGCGTC
V-R_PLIC	ctatagtgagtcgTATTAATTCGCGGGATCG
SaSrtA_F_PLIC	cgactcactatagGGGAATTGTGAGCGGATAAC
SaSrtA_R_PLIC	gaactgcggatgGCTCCATGC

(F: forward primer; R: reverse primer; small letters indicate phosphorothioate deoxynucleotides)

Table. S2 List of primers used for the generation of site-saturation mutagenesis libraries (N includes G or T and M includes A or C).

Primer Name	Sequence 5'-3'
<i>Fw SSM R159</i>	GACAAGTATAN NNK GATGTTAAGCC
<i>Rev SSM R159</i>	GGCTTAACATC MNN TATACTT GTC
<i>Fw SSM D170</i>	GTAGAAGTTC TANNK GAACAAAAAGG
<i>Rev SSM D170</i>	CCTTTTGTTC MNN TAGAACTTCTAC
<i>Fw SSM Q172</i>	GTTCTAGATGA ANNK AAAAGGTAAAG
<i>Rev SSM Q172</i>	CTTTACCTTT MNN TTTTCATCTAGAAC
<i>Fw SSM D186</i>	CATTAATTACTTGTGAT NNK TACAATGAAAAGACAG
<i>Rev SSM D186</i>	CTGTCTTTTCATTGT AMNN ATCACAAGTAATTAATG
<i>Fw SSM K196</i>	GGCGTTTGGGAA NNK CGTAAAATCTTTG
<i>Rev SSM K196</i>	CAAAGATTTTACG MNN TTCCCAAACGCC
<i>Fw SSM D186/K196</i>	TTACTTGTGAT NNK TACAATGAAAAGACAGGCGTTTGGGAA NNK CGTAAAA
<i>Rev SSM D186/K196</i>	AAGATTTTACG MNN TTCCCAAACGCCTGTCTTTTCATTGT AMNN ATCACAAGTAA

Table. S3 Sequencing results of identified variants in Sa-SrtA SSM libraries.

Library	Number in microtiter plate	Mutation (DNA codon)	Substitution (amino acid)
SSM P94	P1F1	CCT/CAT	Pro94His (P94H)
	P1F7	CCT/ACT	Pro94Thr (P94T)
	P1B2	CCT/ACT	Pro94Thr (P94T)
SSM R159	P1F1	AGA/GGG	Arg159Gly (R159G)
	P2B10	AGA/ACT	Arg159Gly (R159T)
	P2E12	AGA/ACG	Arg159Gly (R159T)
SSM D160	P1D9	GAT/AAT	Asp160Asn (D160N)
	P1E10	GAT/AAT	Asp160Asn (D160N)
	P1E11	GAT/AAT	Asp160Asn (D160N)
SSM D165	P1B10	GAT/CAG	Asp165Gln (D165Q)
	P1C11	GAT/CAG	Asp165Gln (D165Q)
	P1H7	GAT/GCG	Asp165Ala (D165A)
SSM D170	P1B8	GAT/TGG	Asp170Trp (D170W)
	P2D5	GAT/TGG	Asp170Trp (D170W)
	P2F3	GAT/GAT	Wild-type
SSM Q172	P2C6	CAA/CAA	Wild-type
	P2E2	CAA/TGG	Gln172Trp (Q172W)
	P2H1	CAA/CTA	Gln172Leu (Q172L)
SSM D186	P1C11	GAT/GGG	Asp186Gly (D186G)
	P1E8	GAT/GGT	Asp186Gly (D186G)
	P2A7	GAT/GGG	Asp186Gly (D186G)
SSM K196	P1C1	AAA/GTG	Lys196Val (K196V)
	P1D9	AAA/GTG	Lys196Val (K196V)
	P2F12	AAA/GTG	Lys196Val (K196V)

Table. S4 List of primers used in the site-directed mutagenesis (N means G or T and M means A or C).

Primer Name	Sequence 5'-3'
<i>Fw SDM R159G</i>	GACAAGTATAGGAGATGTTAAGCCAAC
<i>Rev SDM R159G</i>	GTTGGCTTAACATCTCCTATACTTGTC
<i>Fw SDM R159T</i>	GACAAGTATAACGGATGTTAAGCCAAC
<i>Rev SDM R159T</i>	GTTGGCTTAACATCCGTTATACTTGTC
<i>Fw SDM D165A</i>	GTAAAGCCAACAGCTGTAGAAGTTCTAGATG
<i>Rev SDM D165A</i>	CATCTAGAACTTCTACAGCTGTTGGCTT AAC
<i>Fw SDM D165Q</i>	GTAAAGCCAACACAGGTAGAAGTTCTAGATG
<i>Rev SDM D165Q</i>	CATCTAGAACTTCTACCTGTGTTGGCTTAAC
<i>Fw SDM D170W</i>	GTAGAAGTTCTATGGGAACAAAAAGG
<i>Rev SDM D170W</i>	CCTTTTGTTCCTATAGAACTTCTAC
<i>Fw SDM K196T</i>	GACAGGCGTTTGGGAAACACGTAAAATCTTTGTAG
<i>Rev SDM K196T</i>	CTACAAAGATTTTACGTGTTCCCAAACGCCTGTC
<i>Fw SDM D160N/D165A</i>	GACAAGTATAAGAAATGTTAAGCCAACAGCTGTAGAAGTTCTAGATGAAC
<i>Rev SDM D160N/D165A</i>	GTTTCATCTAGAACTTCTACAGCTGTTGGCTTAACATTTCTTATACTTGTC

Table. S5 Activity of Sa-SrtA variants in absence and presence of DMSO using FRET assay.⁸ The variants selected for further characterization are marked in red.

Sa-SrtA WT / variants	Specific activity in absence of DMSO (slope (RFU/s))	Relative activity in absence of DMSO (variant/WT) ^a	Specific activity in 45% (v/v) DMSO (slope (RFU/s))	Relative activity in 45% (v/v) DMSO (variant /WT) ^b	Resistance (%) in 45% (v/v) DMSO ^c
WT	3.34±0.08	1.00	0.55±0.06	1.00	16.39
M1 (R159G)	2.15±0.03	0.73	0.83±0.04	1.52	33.88
R159T	7.23±0.1	2.16	1.11±0.11	2.02	15.35
D170W	6.88±0.11	2.06	1.16±0.03	2.10	16.86
Q172L	2.80±0.22	0.84	0.67±0.07	1.22	23.93
D186G	10.36±0.48	3.10	1.23±0.09	2.24	11.87
K196V	6.46±0.24	1.90	1.02±0.13	1.85	16.10
M2 (D186G/K196V)	19.31±0.476	5.78	1.66±0.09	3.02	8.61
M4 (P94H/D186G/K196V)	18.72±0.57	5.60	1.47±0.11	2.68	7.85
M5 (R159G/D186G/K196V)	7.76±0.54	2.33	1.21±0.12	2.22	15.62
M6 (R159T/D186G/K196V)	31.32±0.68	9.38	1.68±0.14	3.07	5.42
M7 (D165A/D186G/K196V)	59.72±1.39	17.80	2.53±0.12	4.63	4.23
M3 (D165Q/D186G/K196V)	56.76±1.32	17.01	2.99±0.13	5.46	5.26
M8 (D170W/D186G/K196V)	15.88±0.27	4.76	1.13±0.07	2.06	7.10
rM4 (P94S/D160N/D165A/K196T)	265.39±3.27	79.46	1.01±0.04	1.84	0.37

^aRelative activity was calculated the ratio of specific activity of Sa-SrtA variants in absence of DMSO divided specific activity of Sa-SrtA WT in absence of DMSO

^bRelative activity was calculated the ratio of specific activity of Sa-SrtA variants in 45% (v/v) DMSO divided specific activity of Sa-SrtA WT in 45% (v/v) DMSO

^cResistance in 45% (v/v) DMSO was calculated as the ratio (percentage) of specific activity in 45% (v/v) DMSO divided specific activity in absence of DMSO.

Table. S6 Relative activity of Sa-SrtAs in 45% (v/v) DMSO, 30% (v/v) DMF, 30% (v/v) ethanol and 50% (v/v) methanol. Relative activity was calculated as the ratio of variant's activity in presence of a certain concentration of solvent divided by Sa-SrtA WT's activity in presence of corresponding concentration of solvent.

Sa-SrtA	Relative activity in 45% (v/v) DMSO (Variant/WT)	Relative activity in 30% (v/v) DMF (Variant/WT)	Relative activity in 30% (v/v) Ethanol (Variant/WT)	Relative activity in 50% (v/v) Methanol (Variant/WT)
WT	1.00±0.11	1.00±0.14	1.00±0.07	1.00±0.10
M1 (R159G)	1.52±0.05	1.21±0.07	0.94±0.10	1.07±0.14
M3 (D165Q/D186G/K196G)	5.46±0.04	8.58±0.21	5.99±0.16	5.37±0.14
rM4 (P94S/D160N/D165A/K196T)	1.84±0.04	6.74±0.17	4.00±0.12	5.77±0.10

List of supplementary Figures

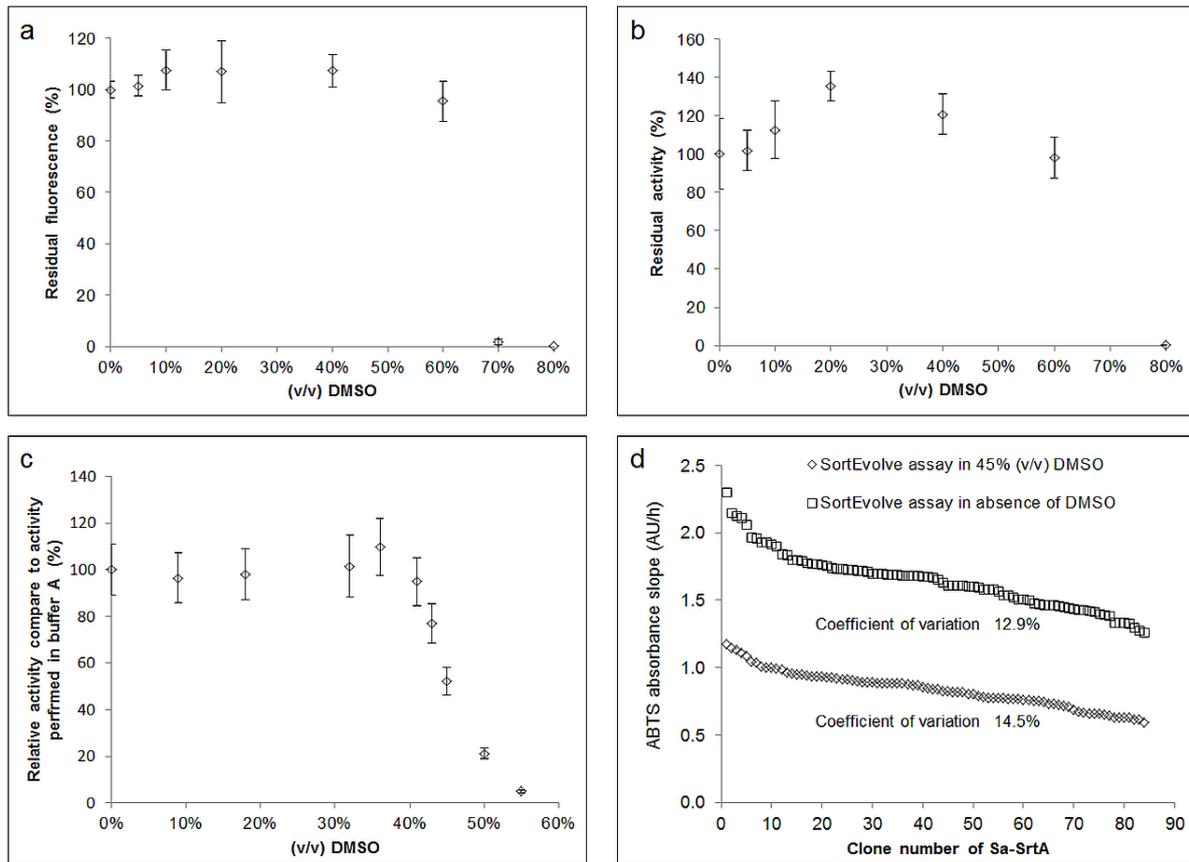


Fig. S1 Optimization of the SortEvolve assay in DMSO co-solvent. a) Stability of GGG-eGFP-LCI in gradient concentration of DMSO; b) Stability of CueO-LPETGGGRR in gradient concentration of DMSO; c) Residual activity of SortEvolve assay performed in gradient concentration of DMSO (A control using pET-28a empty vector lysate instead of Sa-SrtA WT lysate was substrate in all cases); d) Coefficient of variations of SortEvolve assay in absence and in presence of 45% (v/v) DMSO.

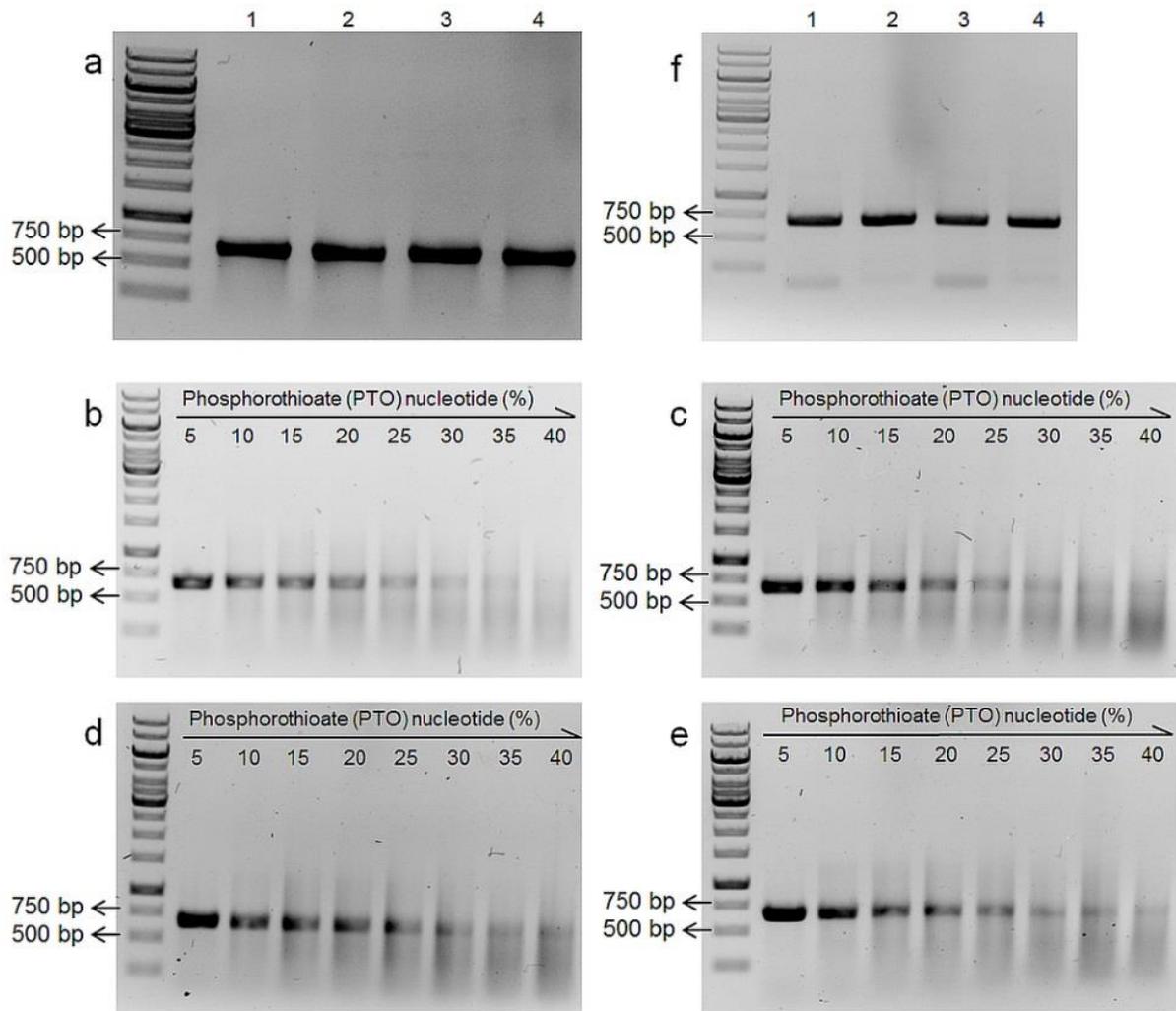


Fig. S2 Agarose gel electrophoretic analysis of SeSaM steps during construction of Sa-SrtA library. Steps were performed as described.⁴ a) Preliminary step 1: amplification PCR to generate template for step 1 and step 3 (expected band size: 0.55 kb). 1: Step 1 forward library template, 2: Step 1 reverse library template, 3: Step 3 forward library template, 4: Step 3 reverse library template. b) – e) Optimization of phosphorothioate deoxynucleotides percentage using gradient concentrations of dATP α S or dGTP α S for A-forward library (b), A-reverse library (c), G-forward library (d), G-reverse library (e). f) Final A and G libraries.

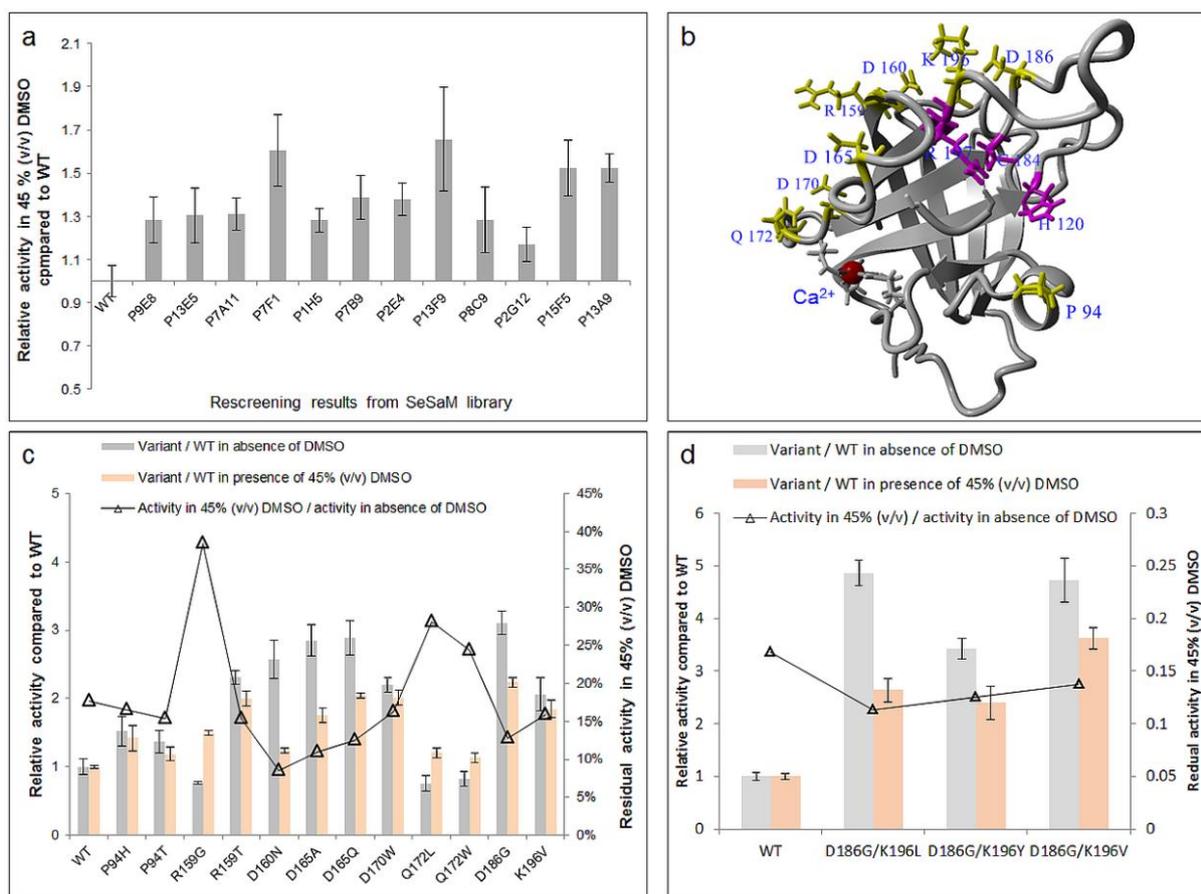


Fig. S3 Directed Sa-SrtA evolution in DMSO co-solvent. a) Rescreening results from random mutagenesis SeSaM library in DMSO co-solvent. Data is shown by ratios of Sa-SrtA variant activity (background subtracted) divided Sa-SrtA WT activity (background subtracted). b) Visualization of the identified positions in Sa-SrtA (PDB 1T2P), identified positions (yellow), active site (magenta) and calcium ion (red) are shown. c) Rescreening results from single-site SSM libraries both in buffer and in 45% (v/v) DMSO co-solvent. Data is shown by ratio of Sa-SrtA variant activity (background subtracted) divided Sa-SrtA WT activity (background subtracted). d) Rescreening results from the SSM D186G/K196V library both in buffer and in DMSO co-solvent. Data is shown by ratio of Sa-SrtA variant activity (background subtracted) divided by Sa-SrtA WT activity (background subtracted).

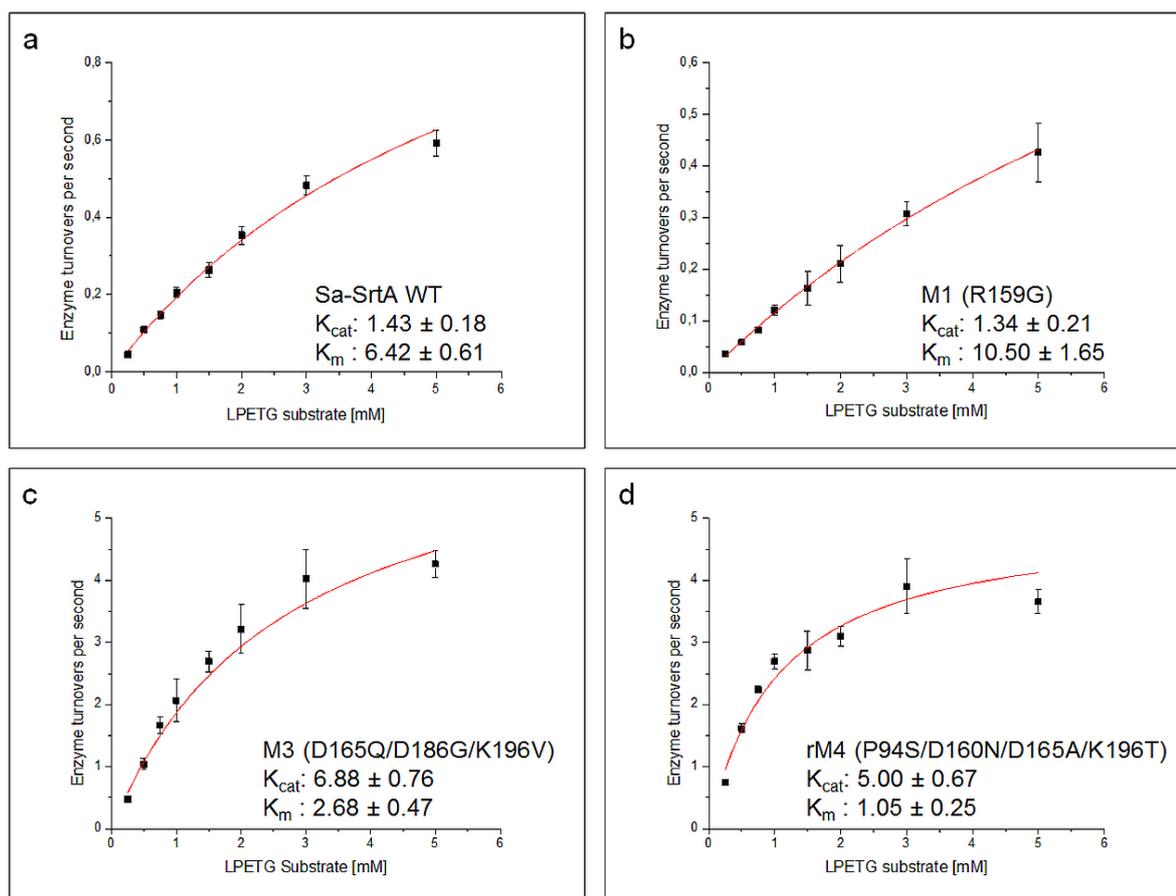


Fig. S4 plots to determine K_m (LPETG) and k_{cat} of Sa-SrtAs in buffer. Plots of Sa-SrtA WT (a), R159G (b), D165Q/D186G/K196V (c) and P94S/D160N/D165A/K196T (d) turnovers upon different Abz-LPETGK-Dpn concentrations in buffer A.

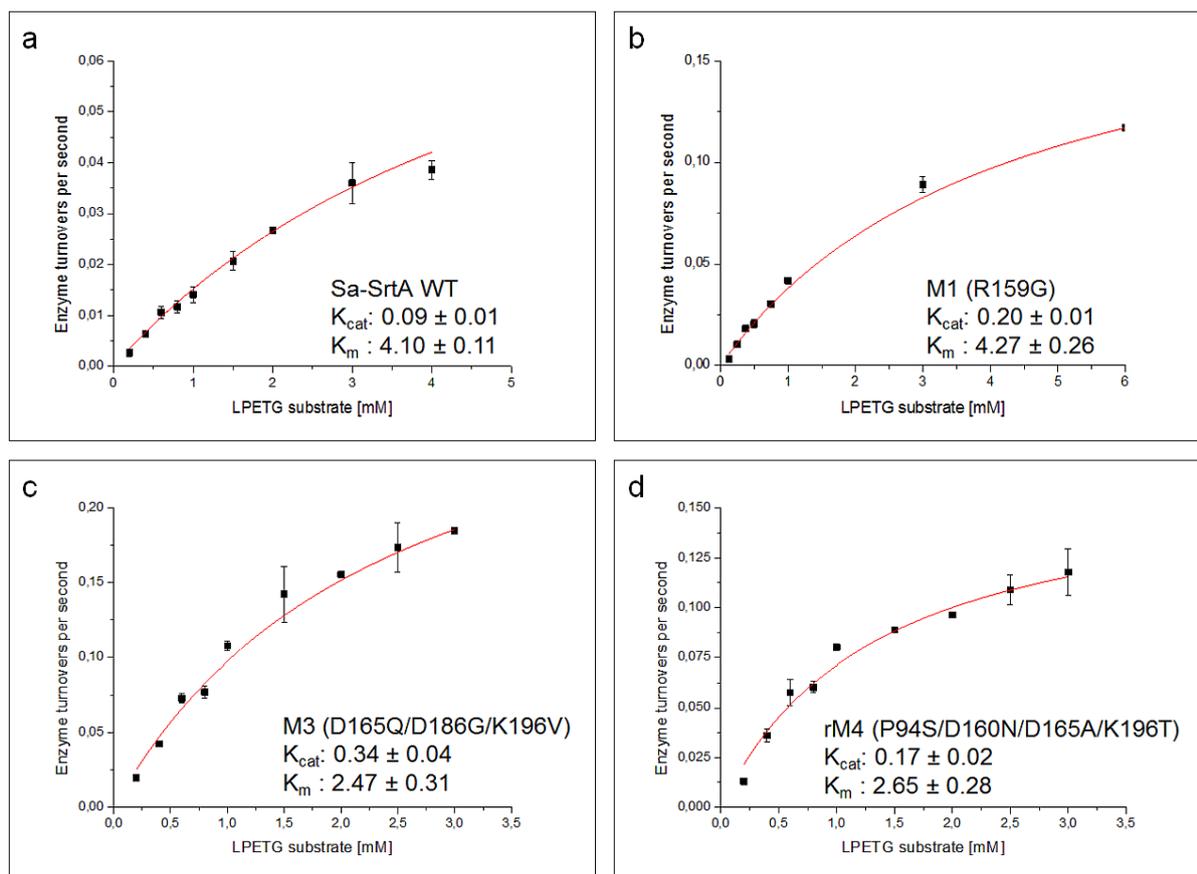


Fig. S5 plots to determine K_m (LPETG) and k_{cat} of Sa-SrtAs in 45% (v/v) DMSO. Plots of Sa-SrtA WT (a), R159G (b), D165Q/D186G/K196V (c) and P94S/D160N/D165A/K196T (d) turnovers upon different Abz-LPETGK-Dpn concentrations in 45% (v/v) DMSO co-solvent.

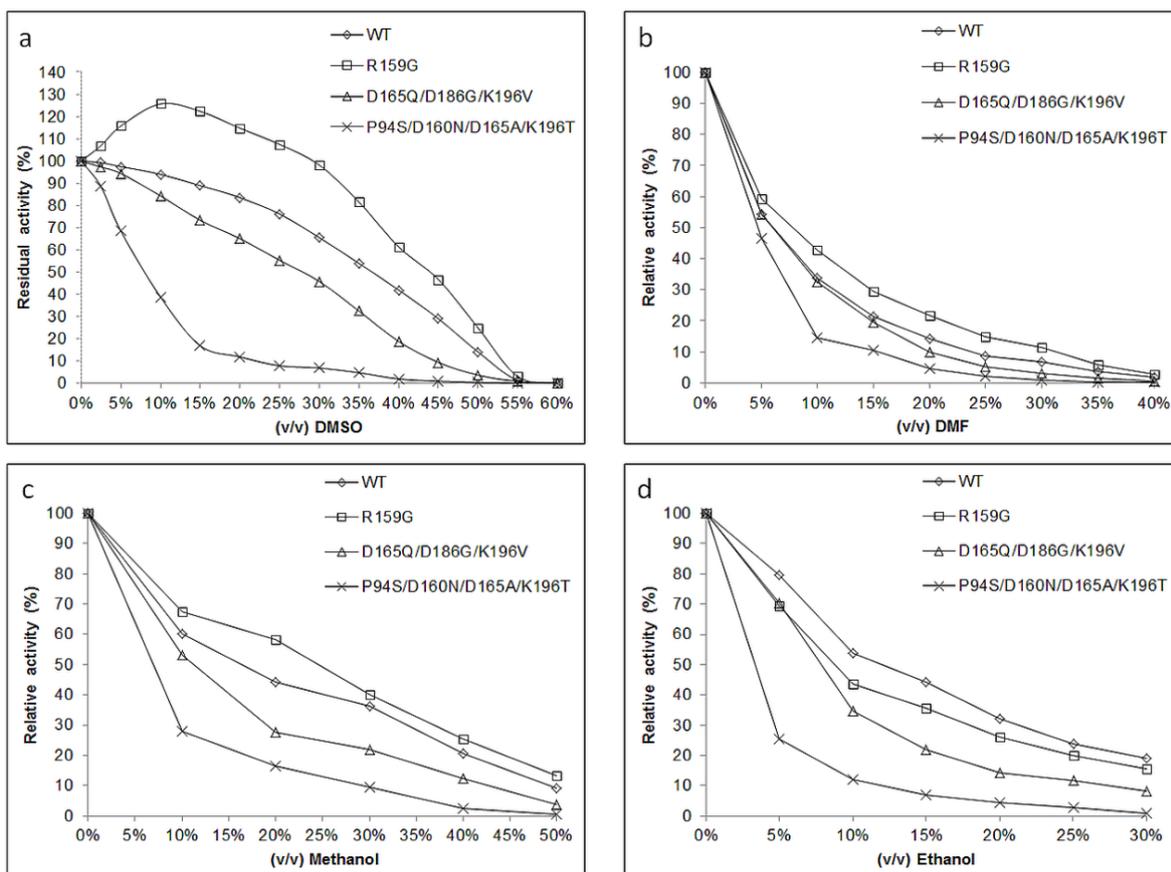


Fig. S6 Activity profiles of Sa-SrtAs in different co-solvents. Activity is recorded by a standard FRET assay using Abz-LPETGK-Dnp as substrate. Residual activity was calculated as the ratio of activity in presence of solvent divided activity in absence of solvent. a) gradient concentration of DMSO; b) gradient concentration of DMF; c) gradient concentration of methanol; d) gradient concentration of ethanol. Sa-SrtA WT and variants R159G (M1), D165Q/D186G/K196V (M3) and P94S/D160N/D165A/K196T were selected for the analysis.

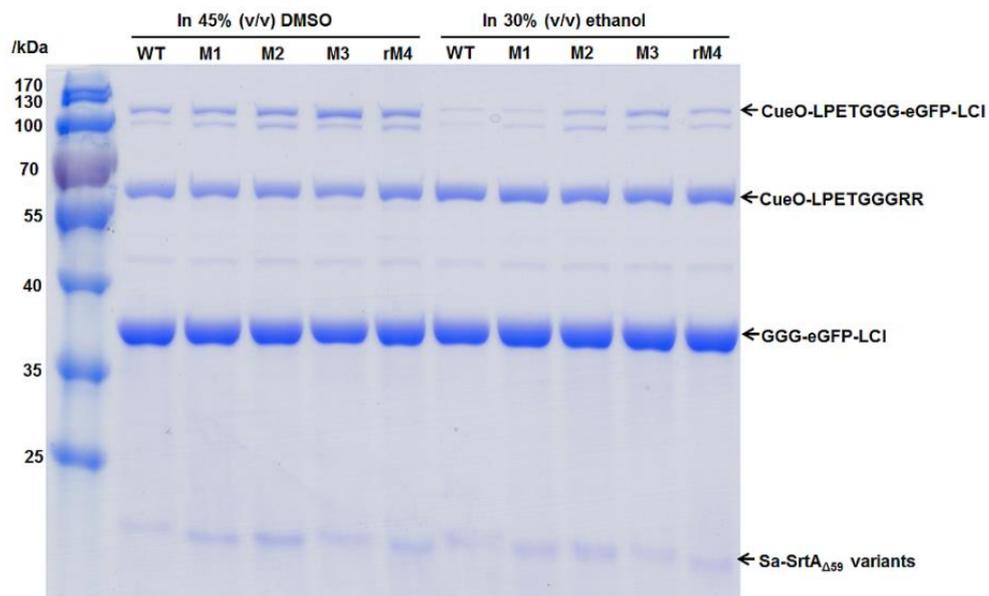


Fig. S7 SDS-PAGE of sortase-mediated protein-protein ligations in 45% (v/v) DMSO or 30% (v/v) ethanol. Reaction mixture contains (500 μ L, 500 μ g/mL purified GGG-eGFP-LCI, 250 μ g/mL purified CueO-LPETGGRR, 30 μ g/mL purified Sa-SrtA, 45% (v/v) DMSO or 30% (v/v) ethanol). Reactions were performed at room temperature for 14 h.

Section 2: Computational study of Sa-SrtA WT and variants in DMSO co-solvent

In silico generation of sortase variants

Structural models of the sortase variants were designed in YASARA Structure version 13.9.8.^{9, 10} using the YASARA-FoldX plugin¹⁰ and by employing the FoldX method.¹¹ The starting coordinates for the FoldX ¹¹ *in silico* mutagenesis were taken from the X-ray structure of the Sa-SrtA WT, chain A (PDB ID: 1T2P for wild type, resolution: 2 Å¹²). A FoldX mutation run including rotamer search, exploring alternative conformations (3 independent runs) were performed during the FoldX energy minimization. Stabilization energy calculations were computed with FoldX version 3.0 Beta¹¹ using standard settings. Calculated stabilization energy ($\Delta\Delta G$) corresponds to the Gibbs free energy changes upon substitution of amino acid residues in unfolded and folded states compared to Sa-SrtA WT. Results shown in **Fig. S8** revealed that the substitution R159G in variant M1 showed a slight destabilizing effect (values less than cutoff 1.0 kcal/mol are considered as stable substitutions). In contrast, the substitutions in variant M3 showed a stabilizing effect of about 0.3 kcal/mol ($\Delta\Delta G < 0$ kcal/mol). Structures of most stable variants (run 1 for M1 and run 3 for M3) were used for further molecular dynamics (MD) simulations studies.

Molecular dynamics (MD) simulations

Molecular dynamics (MD) simulations were used to derive statistical properties of the water and/or co-solvents and analyze the sortase-co-solvent interactions leading to stabilization or destabilization. The main factors in MD simulation studies for describing sortase activity or resistance in DMSO/water mixtures were: (a) DMSO co-solvent-sortase interactions; (b) conformational changes of the sortase structure. We have elucidated the important properties that can differentiate the resistant and more active sortase variants (M1 and M3) from wild-type (WT) based on MD simulation results. MD simulations and analysis were performed using GROMACS 5.1.2 software.¹³⁻¹⁷ The GROMOS96 (53a6) force field¹⁸ was used for the simulations of sortase in water as single solvent system and co-solvent with 45% (v/v) DMSO. Chain A of crystal structures (PDB ID: 1T2P for wild type, resolution: 2 Å¹²) were taken as starting structure for simulations. The protonation states were determined with pKa estimation using PROPKA method using the PDB2PQR server.¹⁹⁻²¹ Hydrogen atoms were added by assuming conventional protonation states of the polar side chains: Lys and Arg, positively charged; Glu and Asp, negatively charged. The His side chains were protonated following the analysis of their environments. The Gln and Asn amide-group rotamers were verified by an inspection of their local interactions. Structures were solvated into a cubic box of SPCE²² water molecules using periodic boundary. Simulation box with volume of 326 nm³ was used. The simulation box was filled with around 9886 water molecules in water system, 1185 DMSO molecules and 6343 water molecules in 45% DMSO system.

Furthermore, in order to neutralize the system, Na⁺ or Cl⁻ ions were added into simulation box. Prior to simulations, energy minimization of the whole system was performed individually using steepest descent minimization algorithm until the maximum force reached to 1000.0 kJ mol⁻¹ nm⁻¹. The electrostatic interactions were calculated by applying the particle mesh Ewald (PME) method.^{23, 24} Short-range electrostatic interactions (rcoulomb) and Van der Waals (rvdwl)

were calculated using a cut-off value 1.0 respectively. Subsequently, system equilibration was performed under an NVT ensemble and NPT ensemble. First, NVT equilibration was conducted at constant temperature of 300 K for 100 ps with time step of 0.002 ps. Initial random velocities were assigned to the atoms of the molecules according to the Maxwell–Boltzmann algorithm at same temperature. Second, NPT equilibration was conducted at constant temperature of 300 K for 100 ps with time step of 2 fs, respectively. The Berendsen thermostat and Parrinello–Rahman pressure coupling were used to keep the system at 300 K, time constant (τ_T) of 0.1 ps and 1 bar pressure, time constant (τ_p) of 2 ps. The production run was carried out in triplicate (run1, run2, run3) using NPT ensemble for 50 ns with time step of 2 fs at constant temperature of 300 K. The coordinates were saved every 200 ps from MD trajectories. All bonds between hydrogen and heavy atoms were constrained with the LINCS algorithm²⁵.

For analyzing the different properties of our system, combined global and local sortase properties analysis was a suitable solution to achieve a deep understanding of molecular interaction between sortase and DMSO co-solvents during MD simulations. Analyses were including root mean square deviation (RMSD) of backbone atoms of the protein with respect to minimized crystal structure,²⁶ root mean square fluctuations (RMSF) per residues,²⁷ radius of gyration (Rg),²⁸ hydrogen-bonding,²⁹ secondary structure change³⁰, spatial distribution function (SDF),³¹ solvent accessible surface areas (SASA)³² of protein. MD trajectories and the structures were analyzed and visualized by using GROMACS analysis tools³³ and VMD 1.9.1 software.³⁴

List of supplementary Figures

Fig. S8

Calculated stabilization energies ($\Delta\Delta G$) in kcal/mol for 3 independent runs of Sa-SrtA variants (M1 and M3) with respect to Sa-SrtA WT by using the FoldX method; $\Delta\Delta G = \Delta G(\text{variant}) - \Delta G(\text{WT})$.

Fig. S9

- a. Model of Sa-SrtA variant M1 (R159G)
- b. Spatial distribution function (SDF) depicting the solvation mechanism of Sa-SrtA M1 from front and back view in water. Density distribution of water molecules are shown in blue. The molecular surface corresponds to the average structure of Sa-SrtA from the 50 ns trajectory, for each solvation system. Two sides of the Sa-SrtA are shown in order to provide a complete view of the surface. Each view has the same orientation as of the Sa-SrtA in (a).
- c. Spatial distribution function (SDF) depicting the solvation mechanism of Sa-SrtA M1 from front and back view in 45% (v/v) DMSO co-solvent. Density distribution of water molecules are shown in blue and DMSO molecules in red. The molecular surface corresponds to the average structure of Sa-SrtA from the 50 ns trajectory, for each solvation system. Two sides of the Sa-SrtA are shown in order to provide a complete view of the surface. Each view has the same orientation as of the Sa-SrtA in (a).

Fig. S10

- a. Model of Sa-SrtA variant M3 (D165Q/D186G/K196V)
- b. Spatial distribution function (SDF) depicting the solvation mechanism of Sa-SrtA M3 from front and back view in water. Density distribution of water molecules are shown in blue. The molecular surface corresponds to the average structure of Sa-SrtA from the 50 ns trajectory, for each solvation system. Two sides of the Sa-SrtA are shown in order to provide a complete view of the surface. Each view has the same orientation as of the Sa-SrtA in (a).
- c. Spatial distribution function (SDF) depicting the solvation mechanism of Sa-SrtA M3 from front and back view in 45% (v/v) DMSO co-solvent. Density distribution of water molecules are shown in blue and DMSO molecules in red. The molecular surface corresponds to the average structure of Sa-SrtA from the 50 ns trajectory, for each solvation system. Two sides of the Sa-SrtA are shown in order to provide a complete view of the surface. Each view has the same orientation as of the Sa-SrtA in (a).

Fig. S11:

a The RMSD of backbone atoms of the Sa-SrtA WT fitted against the minimized X-ray structure shows that the structure of the sortase in the DMSO/water mixture shows less deviation from the crystal structure compared to the simulation in water and the Sa-SrtA WT structure is slightly less flexible in DMSO co-solvent than in water.

b, c To estimate the exposure of sortase to solvents, we calculate the solvent accessible surface area (SASA) and radius of gyration (Rg). Larger average SASA and radius of gyration (Rg) values show the swelling of Sa-SrtA WT in DMSO co-solvent.

d Preferential solvation of sortase through the DMSO or water was analyzed by radial distribution functions (RDF). RDF plot for water molecules (first and second hydration shells) show a decrease in the presence of DMSO co-solvent. As a conclusion also observed by Spatial distribution function (SDF) depicting the solvation mechanism of Sa-SrtA WT (**Fig.2 S9** and **S10**), DMSO strips off bound water molecules from the sortase surface.

e From RMSF per residue it is found that substrate binding residues (catalytic site residues) also show less flexibility in DMSO than in water, this eventually leads to reduced enzymatic activity.

Fig. S12:

a Average number of Hbonds between variant R159G and DMSO molecules is more than Sa-SrtA WT along simulation trajectory.

b Average number of Hbonds between Sa-SrtA WT or variant R159G and water molecules is decreased in DMSO co-solvent compared to water (due to presence of less number of water molecules in the system); Average number of Hbonds between variant R159G and water molecules is slightly decreased compared to Sa-SrtA WT (tendency for increased Hbond interaction with DMSO rather than water molecules).

Fig. S13:

a The RMSD of backbone atoms of the R159G variant fitted against the minimized initial structure shows that the structure of the sortase in the DMSO/water mixture shows a comparable deviations from the structure compared to the simulation in water. Comparison of RMSD of R159G and Sa-SrtA WT shows that R159G is more flexible in DMSO co-solvent than Sa-SrtA WT in DMSO.

b, c Larger average SASA and radius of gyration (Rg) values for R159G compared to Sa-SrtA WT show more swelling in DMSO co-solvent.

d From RMSF per residue of Sa-SrtA WT, it is found that substrate binding residues (catalytic site residues) show less flexibility in DMSO than in water, this eventually leads to less enzymatic activity. From comparison of RMSF per residue of Sa-SrtA WT and R159G, it is clear that substrate binding site residues (highlighted by an arrow) show higher flexibility in R159G compared to Sa-SrtA WT.

Fig. S14:

a Average number of Hbonds between variant M3 and DMSO molecules is similar to Sa-SrtA WT along simulation trajectory.

b Average number of Hbonds between Sa-SrtA WT or variant M3 and water molecules is decreased in DMSO co-solvent compared to water (due to presence of less number of water molecules in the system); Average number of Hbonds between variant M3 and water molecules is slightly decreased compared to Sa-SrtA WT (tendency for increased Hbond interaction with DMSO rather than water molecules).

Fig. S15:

a The RMSD of backbone atoms of the M3 variant fitted against the minimized initial structure shows that the structure of the sortase in the DMSO/water mixture shows a comparable deviations from the structure compared to the simulation in water. Comparison of RMSD of M1 and Sa-SrtA WT shows that M3 flexibility in DMSO co-solvent is slightly increased compared to Sa-SrtA WT in DMSO.

b, c Slightly larger average SASA and radius of gyration (Rg) values for M3 variant compared to Sa-SrtA WT show swelling in DMSO co-solvent. Interestingly, comparison of M3 and Sa-SrtA WT shows a comparable size.

d From comparison of RMSF per residue of Sa-SrtA WT and M3, it is clear that substrate binding site residues (highlighted by an arrow) show higher flexibility in M3 compared to Sa-SrtA WT.

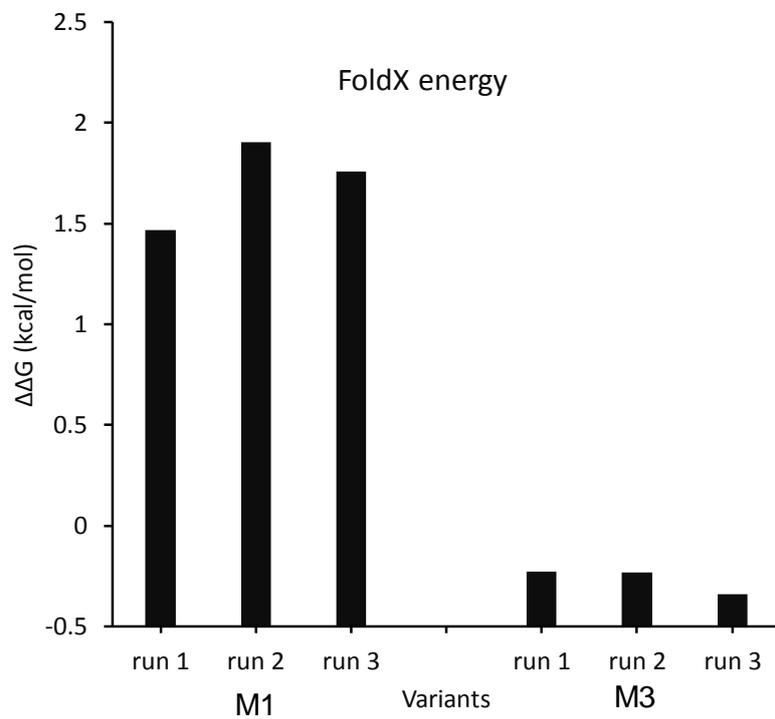


Fig. S8 Calculated stabilization energies ($\Delta\Delta G$) in kcal/mol for 3 independent runs of Sa-SrtA variants (M1 and M3) with respect to Sa-SrtA WT by using the FoldX method; $\Delta\Delta G = \Delta G(\text{variant}) - \Delta G(\text{WT})$.

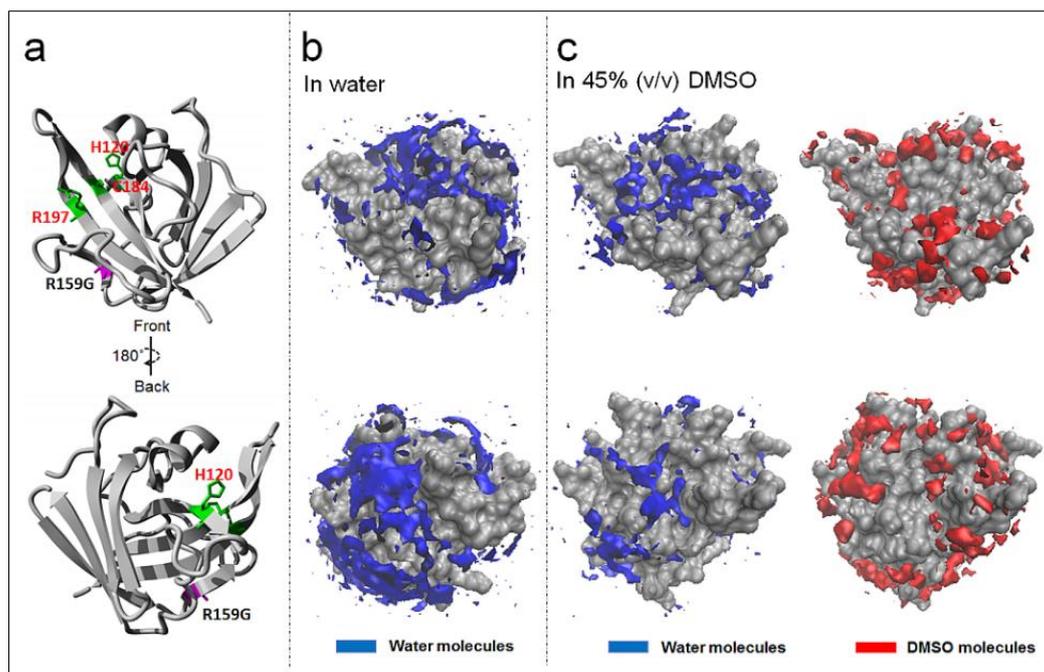


Fig. S9 (a) Model of Sa-SrtA variant M1 (R159G); Spatial distribution function (SDF) depicting the solvation mechanism of Sa-SrtA from front and back view in (b) water and (c) in 45% (v/v) DMSO co-solvent. Density distribution of water molecules are shown in blue and DMSO molecules in red. The molecular surface corresponds to the average structure of Sa-SrtA M1 from the 50 ns trajectory, for each solvation system. Two sides of the Sa-SrtA M1 are shown in order to provide a complete view of the surface. Each view has the same orientation as of the Sa-SrtA M1 in (a).

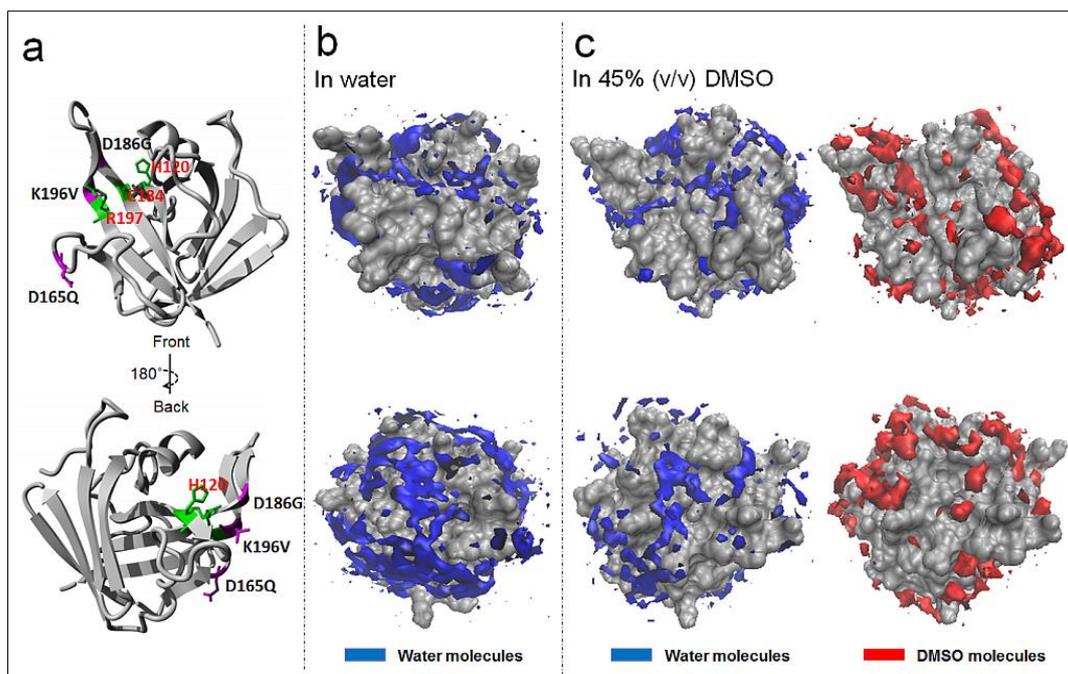


Fig. S10 (a) Model of Sa-SrtA variant M3 (D165Q/D186G/K196V); Spatial distribution function (SDF) depicting the solvation mechanism of Sa-SrtA M3 from front and back view in (b) water and (c) in 45% (v/v) DMSO co-solvent. Density distribution of water molecules are shown in blue and DMSO molecules in red. The molecular surface corresponds to the average structure of Sa-SrtA M3 from the 50 ns trajectory, for each solvation system. Two sides of the Sa-SrtA M3 are shown in order to provide a complete view of the surface. Each view has the same orientation as of the Sa-SrtA M3 in (a).

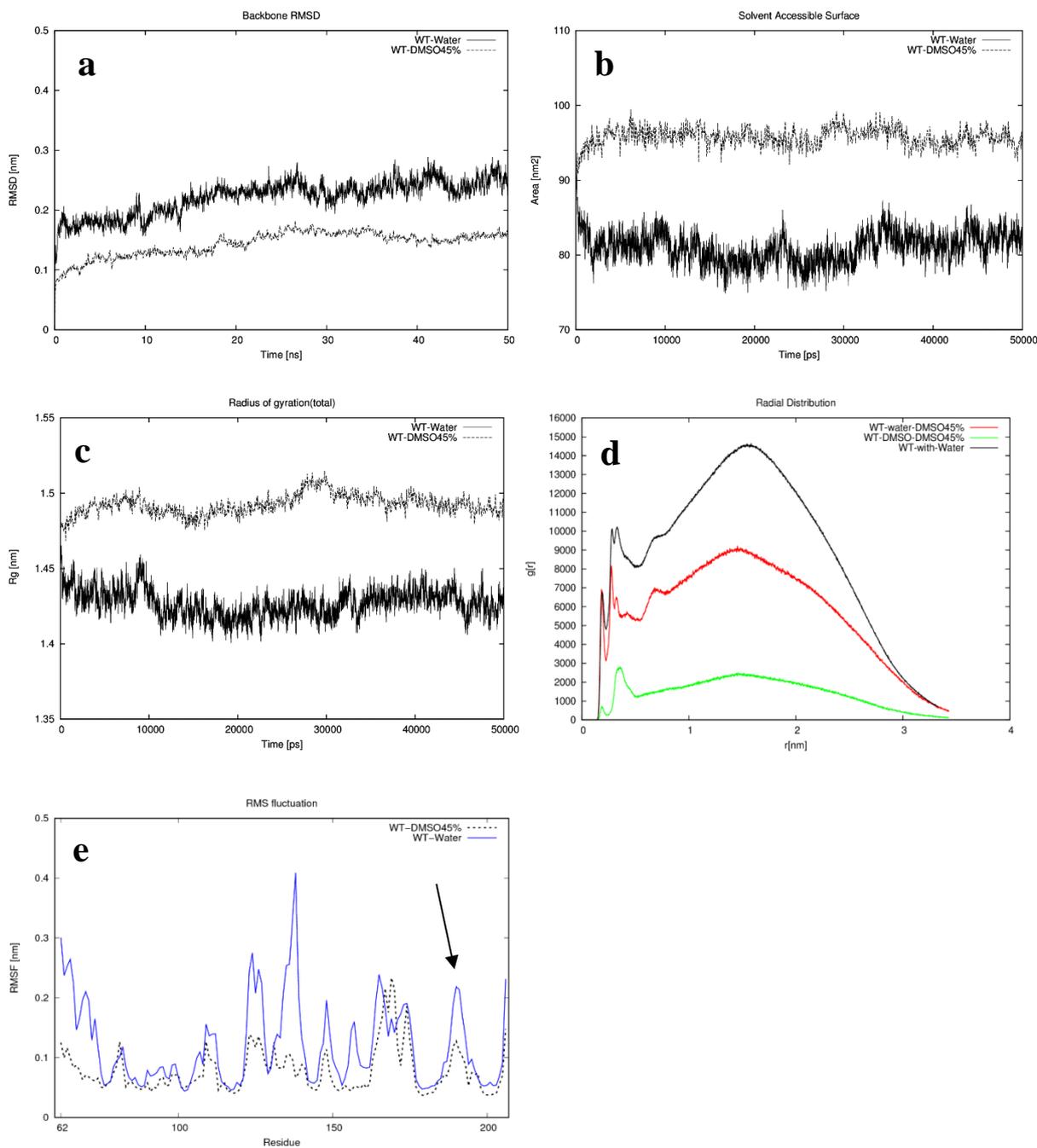


Fig. S11 Analysis of Sa-SrtA WT MD simulation trajectory in water and DMSO co-solvent; **(a)** RMSD of backbone atoms; **(b)** Solvent accessible surface area (SASA); **(c)** radius of gyration (Rg) of sortase; **(d)** radial distribution functions (RDF) of water and DMSO molecules in the surface of Sa-SrtA WT. **(e)** root mean square fluctuation (RMSF) per residues of sortase (catalytic residues are R197, H120, C184). Arrow highlights the substrate binding site region.

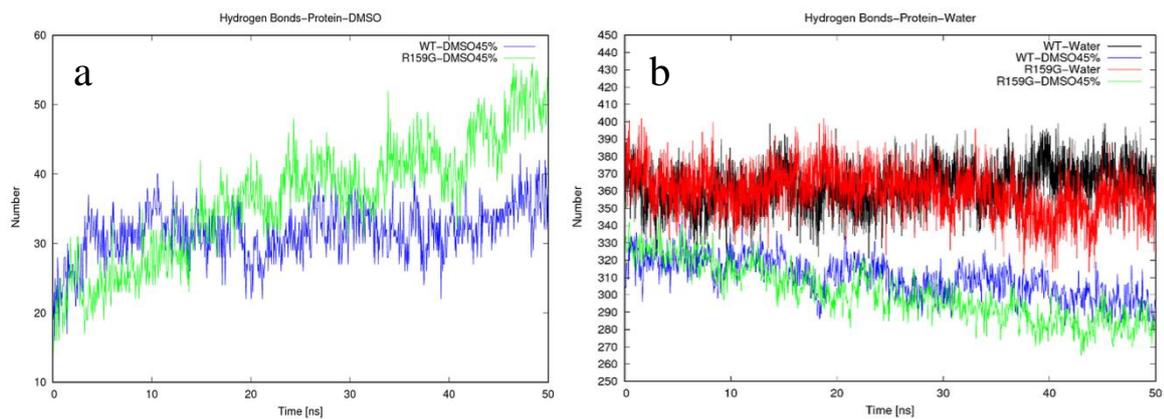


Fig. S12 Analysis of R159G variant MD simulation trajectory; **(a)** number of hydrogen bond between Sa-SrtAs and DMSO, **(b)** number of hydrogen bond between Sa-SrtAs and water.

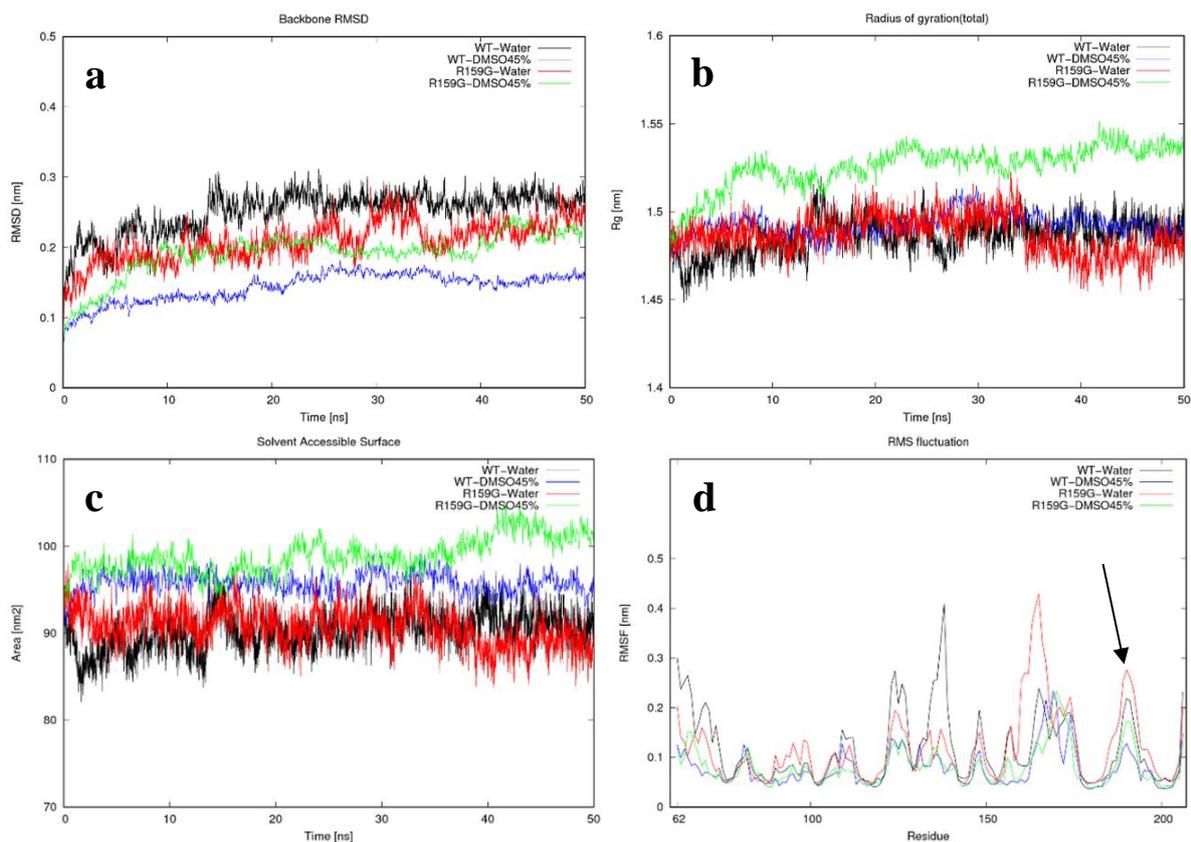


Fig. S13 Analysis of **M1** (R159G) MD simulation trajectory; **(a)** RMSD of back bone atoms for sortase; **(b)** radius of gyration of sortase **(c)** Solvent accessible surface area (SASA), and **(d)** root mean square fluctuation (RMSF) per residues of sortase (catalytic residues are R197, H120, C184). Arrow highlights the substrate binding site region.

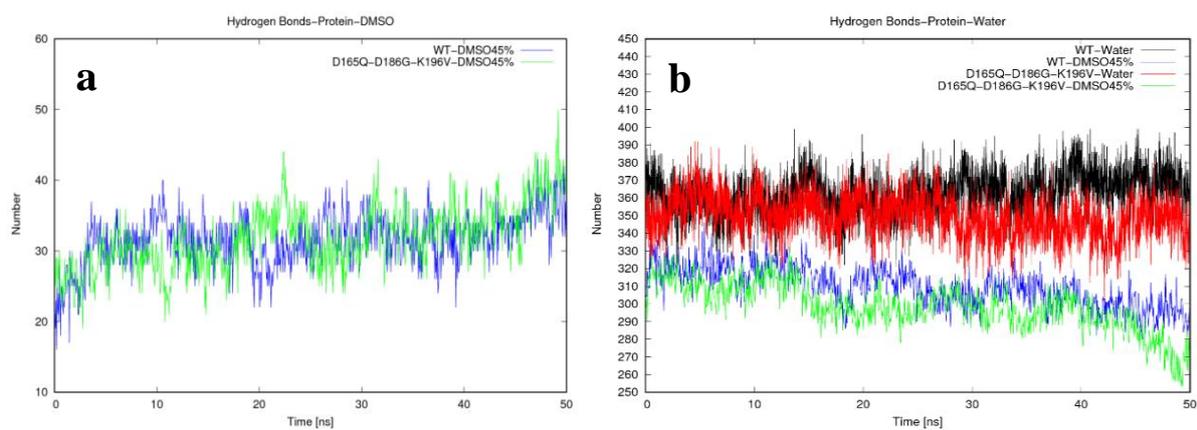


Fig. S14 Analysis of M3 (D165Q/D186G/K196V) variant MD simulation trajectory; (a) number of hydrogen bond between sortase and DMSO, (b) number of hydrogen bond between sortase and water.

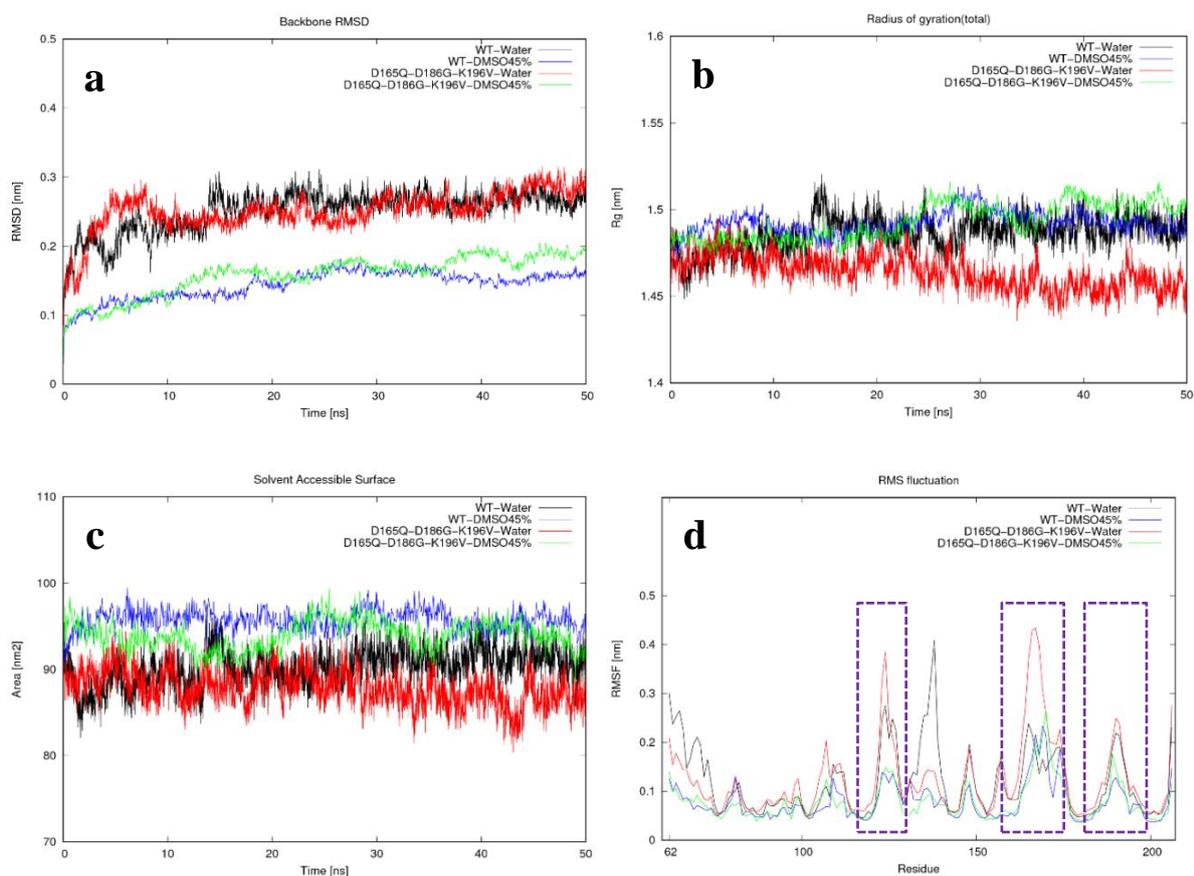


Fig. S15 Analysis of M3 variant (D165Q/D186G/K196V) MD simulation trajectory; **(a)** RMSD of back bone atoms for sortase; **(b)** radius of gyration of sortase **(c)** Solvent accessible surface area (SASA), and **(d)** root mean square fluctuation (RMSF) per residues of Sa-SrtA M3 (catalytic residues are H120, C184 and R197). The three dashed boxes show three regions that are located in the binding site and contain the catalytic residues.

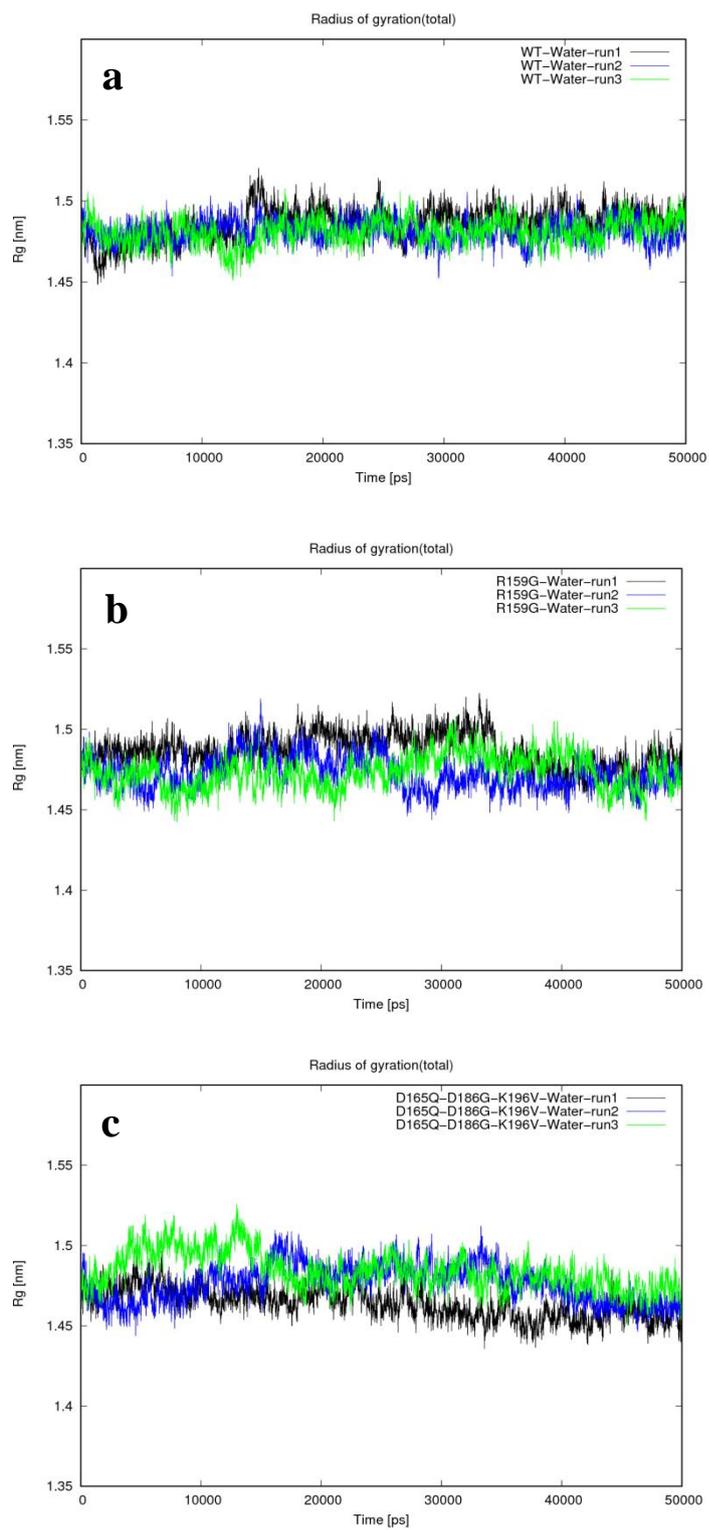


Fig. S16 Analysis of radius of gyration of Sa-SrtA WT (**a**), Sa-SrtA variants M1 (**b**) and M3 (**c**) for three independent MD simulation trajectories (run1, run2, run3).

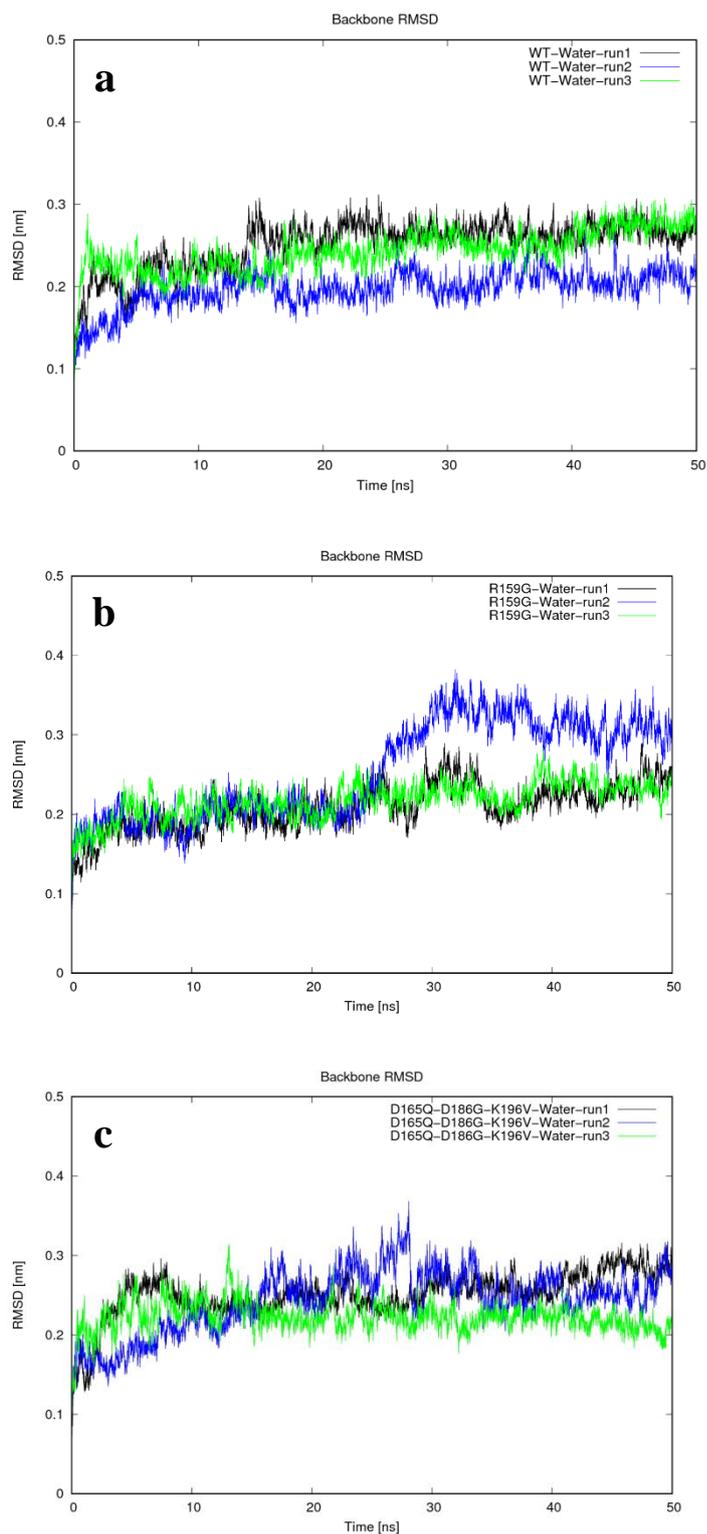


Fig. S17 Analysis of RMSD of back bone atoms of Sa-SrtA WT (a), Sa-SrtA variants M1 (b) and M3 (c) for three independent MD simulation trajectories (run1, run2, run3).

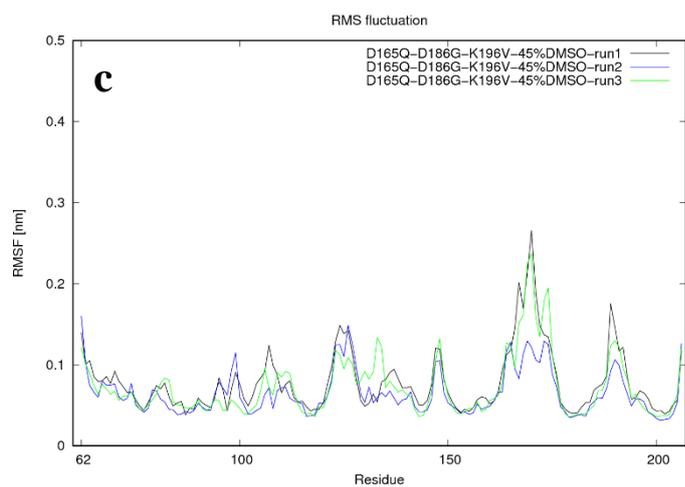
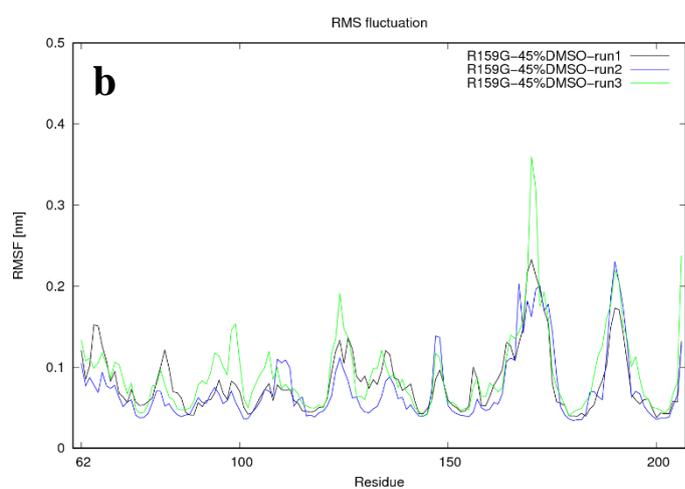
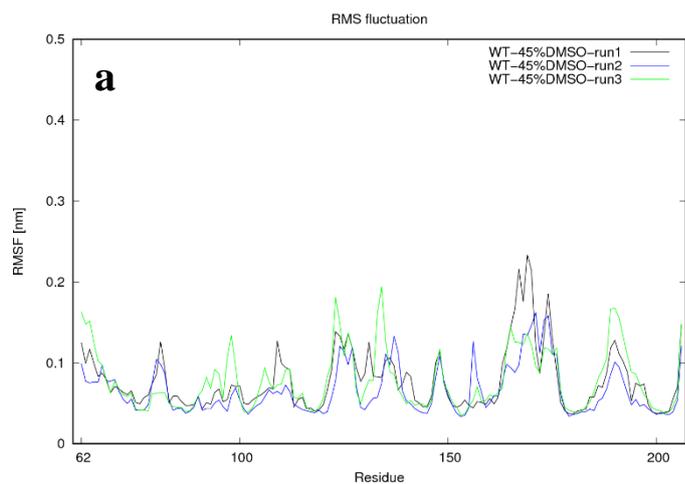


Fig. S18 Analysis of root mean square fluctuation (RMSF) per residues of sortase of Sa-SrtA WT (**a**), Sa-SrtA variants M1 (**b**) and M3 (**c**) for three independent MD simulation trajectories (run1, run2, run3). Catalytic residues are R197, H120, C184.

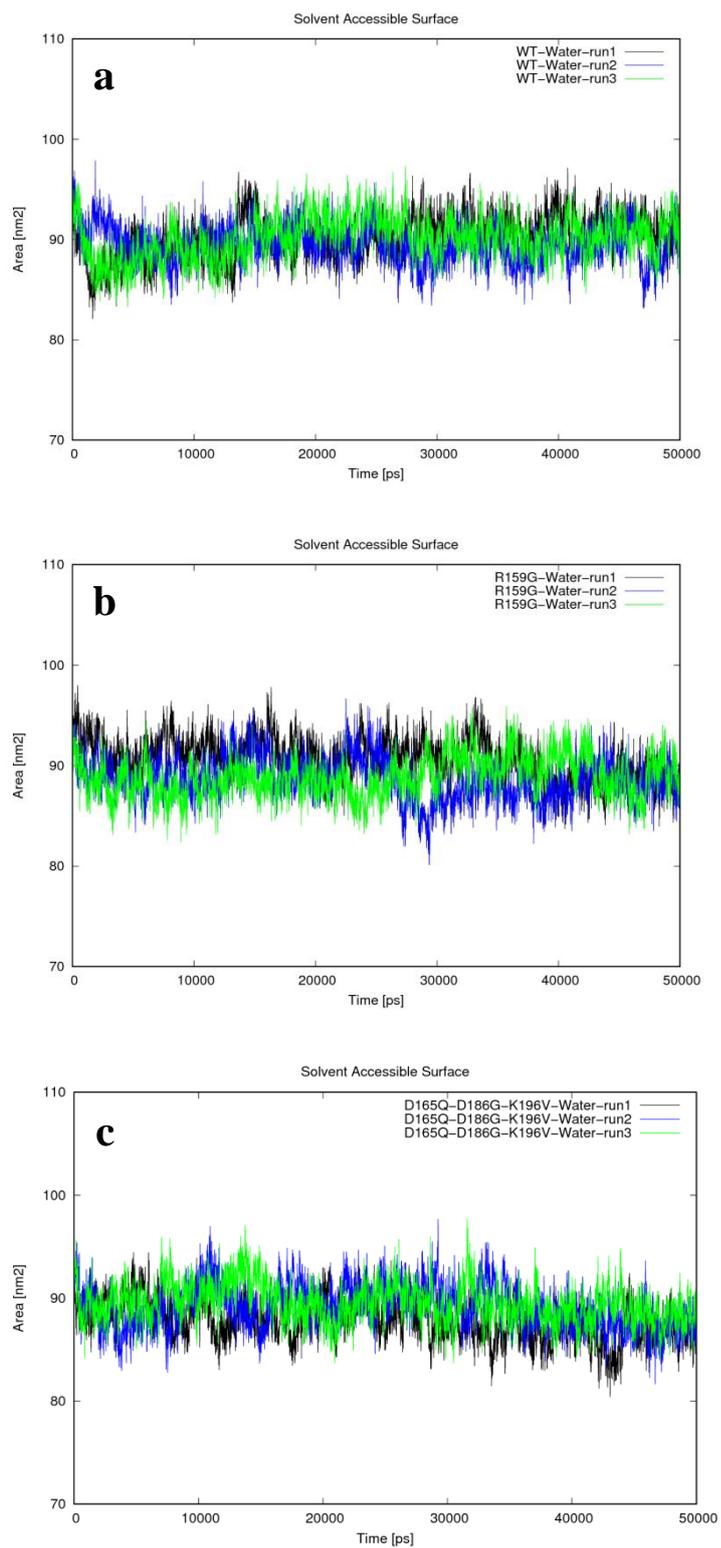


Fig. S19 Analysis of solvent accessible surface area (SASA) of Sa-SrtA WT (a), Sa-SrtA variants M1 (b) and M3 (c) for three independent MD simulation trajectories (run1, run2, run3).

Section 3: Application of engineered sortase A variants for peptide semi-synthesis in organic co-solvents

Materials

Fluorescent peptides Abz-LPETGK-Dnp-NH₂ (97.8%), Abz-LPETGGG-COOH (97.2%) and antiviral peptide AVP0683 purchased from Bachem (Bubendorf, Germany). Tyramine (99%), Trifluoromethyl-benzylamine (4-TFB amine, 99%) and O-(2-Aminoethyl) polyethylene glycol (PEG amine, M_p=5000 da, 99%) were purchased from Sigma-Aldrich (Hamburg, Germany).

Experimental part

Sostase-mediated ligation of hydrophobic antiviral peptide AVP 0683 and Abz-LPETGK-Dnp in DMSO or DMF co-solvent

The solubility of Abz-LPETGK-Dnp and peptide AVP0683 in water, buffer A (buffer A: 5 mM CaCl₂, 150 mM NaCl, 50 mM Tris-HCl buffer, pH 7.5) and DMSO / DMF co-solvents (organic solvents both mix with buffer A) were investigated (**Fig. S20a**). The mixture of 1 mM Abz-LPETGK-Dnp and 3 mM peptide AVP0683 (sequence: GGHRRYFTFGGGYVYF) did not dissolve neither in pure water nor in buffer A. The mixture of 1 mM Abz-LPETGK-Dnp and 3 mM peptide AVP0683 dissolved in 45% (v/v) DMSO and 30% (v/v) DMF.

The ligation of Abz-LPETGK-Dnp and peptide AVP0683 catalyzed by Sa-SrtAs in 45% (v/v) DMSO and 30% (v/v) DMF was performed. In short, reaction solution 1 (100 μL, 1 mM Abz-LPETGK-Dnp, 3 mM peptide AVP0683, 45% (v/v) DMSO) and reaction solution 2 (100 μL, 1 mM Abz-LPETGK-Dnp, 3 mM peptide AVP0683, 30% (v/v) DMF) were incubated in black F-bottom 96-well PS-MTP. Reactions were initiated by adding 10 μM purified Sa-SrtA. Fluorescence was constantly recorded ($\lambda_{exc} = 320$ nm; $\lambda_{em} = 420$ nm, gain = 100, Tecan infinite 1000Pro platereader). After measurement, samples from MTP was subsequently heated at 90°C for 10 min. Samples were stored in 4°C either for subsequently SDS-PAGE, HPLC or MALDI-TOF MS analysis (matrix-assisted laser desorption ionization time-of-flight mass spectroscopy). Three microliter was added into 57 μL water (20-fold dilution) and subsequently mixed with 20 μL 4x SDS loading buffer. The mixture was heated at 95°C for 5 min. Three microliters of the samples were loaded for SDS-PAGE analysis. The SDS-PAGE gel is shown in **Fig. S20c**. In order to separate the fusion product Abz-LPETG-AVP0683 (theoretical molecular weight 2446.6 Dalton) from the peptide substrate AVP0683 (molecular weight 1884.7 Dalton), a modified 25% tricine acrylamide gels³⁵ was used. Electrophoresis was performed with a constant current at 6 mA for 18 h.

HPLC of the reaction mixture was performed based on the protocol as aforementioned. In brief, 20 μL reaction mixture was then diluted 5-times (final volume: 100 μL) with pure water. Twenty microliter of the diluted sample was injected into a reversed-phase C18 HPLC column (4.6x150 mM, 5 μM, Macherey-Nagel, Düren, Germany) and chromatographed using a gradient of 10 to 40% acetonitrile with 0.1% TFA (trifluoroacetic acid) in 0.1% aqueous TFA over 20 minutes. Dnp containing peaks were detected at 355 nm. The yield of product GK-Dnp was calculated by integrating the area under HPLC trace (**Fig. S21**).

MALDI-TOF MS results are shown in **Fig. S22**. Molecule peaks at 2449.3 Dalton (in 45% (v/v) DMSO) and 2451.2 Dalton (in 30% (v/v) DMF) were detected (**Fig. S22b** and **S22c**). Slight differences between theoretical and recorded molecular weight were observed. These might be explained that samples were prepared from buffer A (5 mM CaCl₂, 150 mM NaCl, 50 mM Tris-HCl buffer, pH 7.5) with co-solvents instead of pure water.

Sortase-mediated ligation of hydrophobic amines (tyramine or 4-(Trifluoromethyl)-benzylamine (4-TFB amine)) and Abz-LPETGK-Dnp in DMSO co-solvent

The ligation of Abz-LPETGK-Dnp and tyramine catalyzed by Sa-SrtA WT or M3 in 45% (v/v) DMSO was performed. In short, reaction 1 (100 μ L, 10 mM tyramine ($M_w = 137.18$ da, solubility in pure water $\cong 50$ mM), 1 mM Abz-LPETGK-Dnp, 20 μ M Sa-SrtA, 45% (v/v) DMSO) and reaction 2 (100 μ L, 10 mM 4-(Trifluoromethyl)-benzylamine (4-TFB amine, $M_w = 175.15$ da, solubility in pure water $\cong 5$ mM), 1 mM Abz-LPETGK-Dnp, 20 μ M Sa-SrtA, 45% (v/v) DMSO) were performed. Fluorescence was constantly recorded ($\lambda_{exc} = 320$ nm; $\lambda_{em} = 420$ nm, gain = 100, Tecan infinite 1000Pro plate reader). Reaction was subsequently quenched with HCl (100 μ L, 500 mM). Quenched sample (2 μ L) was analyzed by ultra-performance liquid chromatography mass spectrum (UPLC-MS, Waters ACQUITY UPLC*). Samples were chromatographed using a gradient of 10 to 90% acetonitrile with 0.1% TFA (trifluoroacetic acid) in 0.1% aqueous TFA over 15 minutes. UPLC trace and mass spectrums of expected ligation products were given in **Fig S23** and **S24**. Absorbance of samples was monitored at 355 nm. The theoretical molecular weight of generated conjugates Abz-LPET-tyramine and Abz-LPET-4-TFB are 697.38 and 736.32 da, respectively.

Sortase-mediated PEGylation of hydrophobic peptide Abz-LPETGK-Dnp in DMSO co-solvent

The ligation of Abz-LPETGK-Dnp and O-(2-Aminoethyl) polyethylene glycol (PEG amine, $M_p = 5000$ da) catalyzed by Sa-SrtA WT and M3 in 45% (v/v) DMSO was performed. In short, reaction (100 μ L, 10 mM PEG amine ($M_p = 5000$ da), 1 mM Abz-LPETGK-Dnp, 20 μ M Sa-SrtA, 45% (v/v) DMSO) was performed. Fluorescence was constantly recorded ($\lambda_{exc} = 320$ nm; $\lambda_{em} = 420$ nm, gain = 100, Tecan infinite 1000Pro plate reader). Reaction was quenched with HCl (100 μ L, 500 mM) and aliquots (1 μ L) were analyzed by UPLC as aforementioned. UPLC trace is given in **Fig. S25**. MALDI-TOF analysis of the quenched reaction mixture is showed in (**Fig. S26**).

List of supplementary Table

Table. S7 Molecular weight and solubility of primary amine donors in this study

Primary amines donor	Molecular weight (da)	Solubility in pure water (mM)	Conjugate with Abz-LPETGK-dNP	Theoretical molecular weight of conjugate (da)
AVP 0683	1884.7	< 0.3	Abz-LPET-AVP 0683	2446.6
Tyramine	137.18	< 50	Abz-LPET-tyramine	697.38
4-TFB amine	175.15	< 5	Abz-LPET-4-TFB	736.32
PEG-amine	5000 (mean value)	soluble	Abz-LPET-PEG	5545 (mean value)

List of supplementary Figures

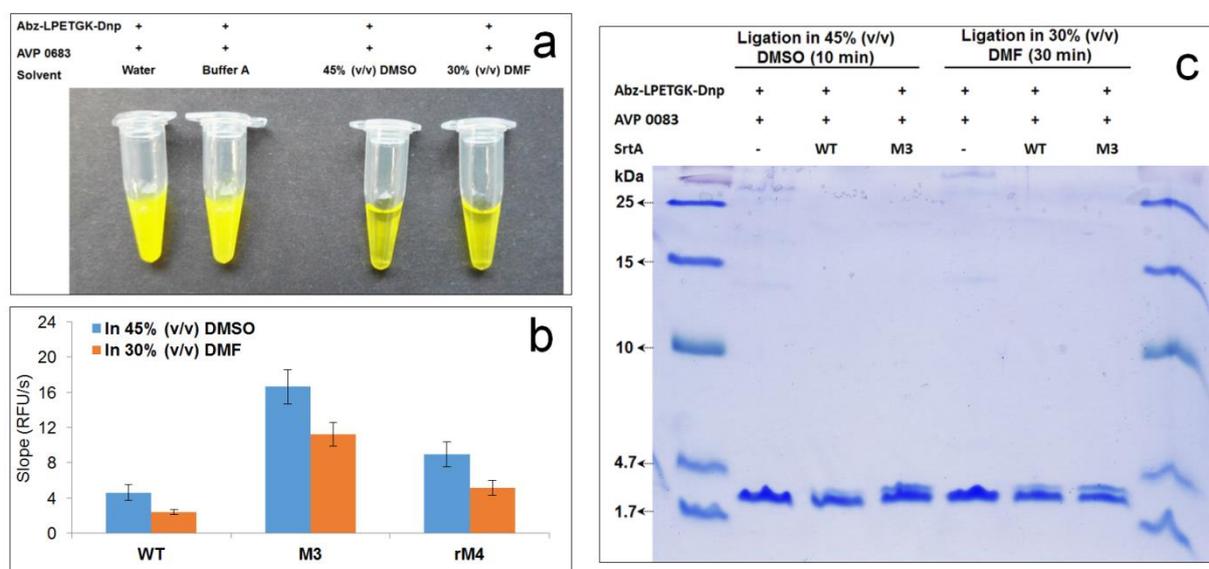


Fig. S20 a) Solubility comparison of 1 mM Abz-LPETGK-Dnp and 3 mM peptide AVP0683 in different solvents. The two co-solvents 45% (v/v) DMSO and 30% (v/v) DMF were dissolved in buffer A. b) Activity of Sa-SrtAs (WT, M3, and rM4) in 45% (v/v) DMSO and 30% (v/v) DMF determined by FRET assay which using peptide AVP0683 and Abz-LPETGK-Dnp as substrates. c) Tricine SDS-PAGE analysis of sortase-mediated ligation (fusion of peptide AVP0683 and Abz-LPETGK-Dnp) in 45% (v/v) DMSO or 30% (v/v) DMF co-solvents.

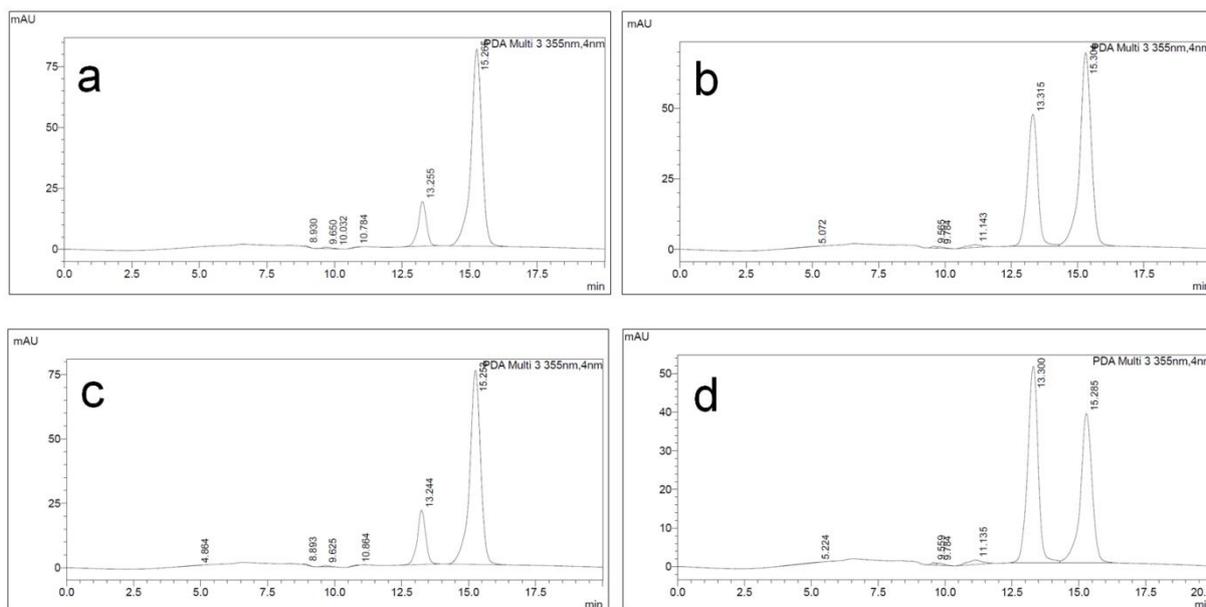


Fig. S21 HPLC trace of reaction mixtures: a) Reaction mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 3 mM peptide AVP0683, 45% (v/v) DMSO, 10 μ M Sa-SrtA WT); b) Reaction mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 3 mM peptide AVP0683, 45% (v/v) DMSO, 10 μ M Sa-SrtA M3); c) Reaction mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 3 mM peptide AVP0683, 30% (v/v) DMF, 10 μ M Sa-SrtA WT); Reaction mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 3 mM peptide AVP0683, 30% (v/v) DMF, 10 μ M Sa-SrtA M3). Absorbance peak at around 13.3 and 15.3 min are generated product GK-Dnp and substrate Abz-LPETGK-Dnp, respectively. Conversions of Abz-LPETGK-Dnp to Abz-LPETG-AVP0683 in (a), (b), (c), and (d) are 13.7 %, 41.9%, 13.3% and 61.8%, respectively. Yield of Abz-LPETGK-Dnp to Abz-LPETG-AVP0683 in (a), (b), (c), and (d) are 0.137, 0.419, 0.133 and 0.618 mM, respectively.

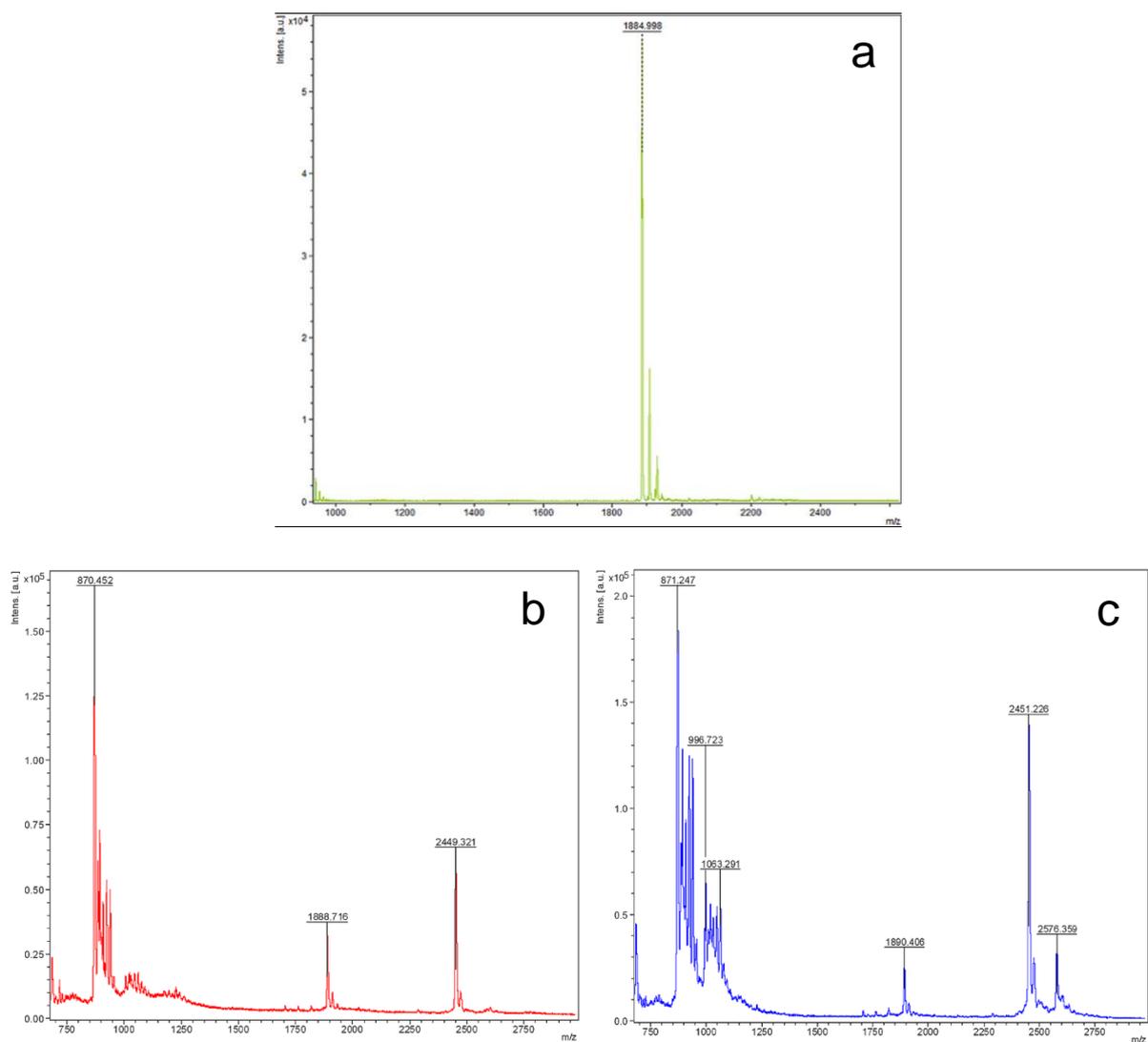


Fig. S22 MALDI-TOF MS of (a) AVP0683, (b) reaction mixture of M3 (10 min) in 45% (v/v) DMSO, and (c) reaction mixture of M3 (30 min) in 30% (v/v) DMF. The theoretical molecular weight of AVP0683 is 1884.7 Da (Dalton) and the theoretical molecular weight of Abz-LPET-AVP0683 is 2446.6 Da (Dalton).

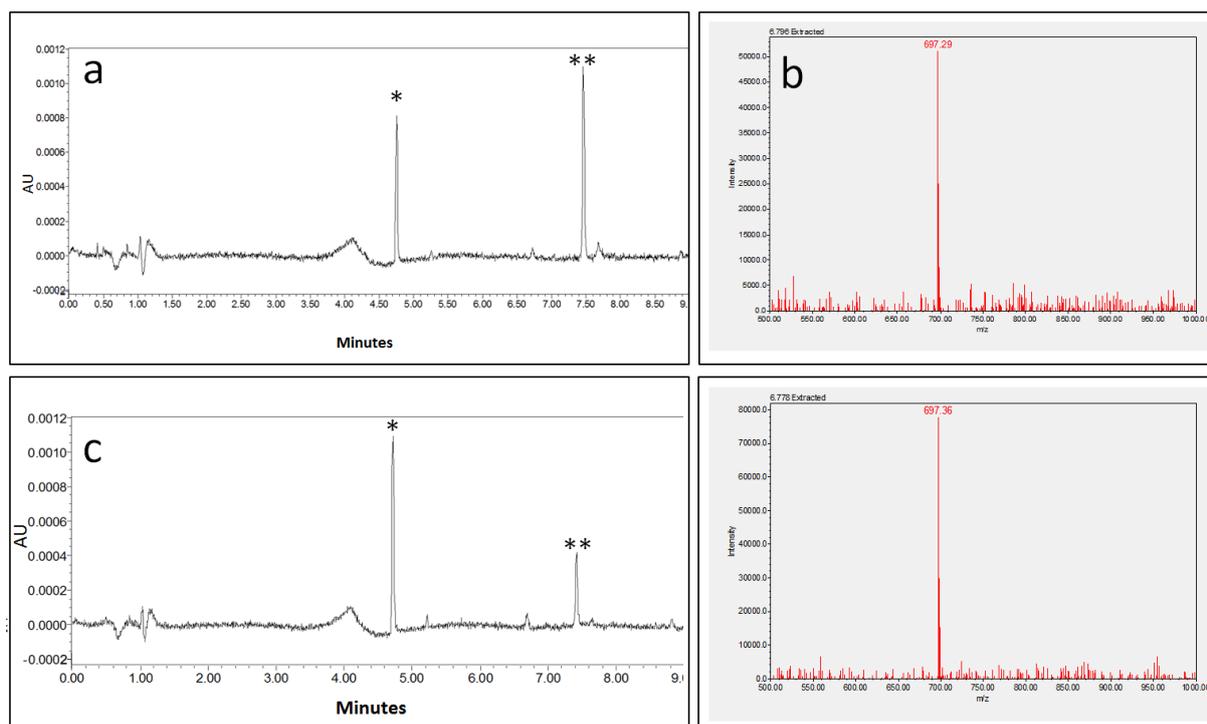


Fig. S23. UPLC-MS of sortase-mediated ligation of tyramine and Abz-LPETGK-Dnp: a) UPLC trace ligation mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 10 mM tyramine, 45% (v/v) DMSO, 20 μ M Sa-SrtA WT); b) Mass spectrum of ligation mixture in (a). The theoretical molecular weight of conjugate (Abz-LPET-tyramine) is 697.38 da. c) UPLC trace ligation mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 10 mM tyramine, 45% (v/v) DMSO, 20 μ M Sa-SrtA M3); d) Mass spectrum of ligation mixture in (c). Absorbance was monitored at 355 nm. Peak* is the generated product GK-Dnp. Peak** is the substrate Abz-LPETGK-Dnp. Conversions of Abz-LPETGK-Dnp to Abz-LPETG-tyramine in (a) and (c) are 39.2 %, 79.6%, respectively. Yield of Abz-LPETGK-Dnp to Abz-LPETG-tyramine in (a) and (c) are 0.392 and 0.796 mM, respectively.

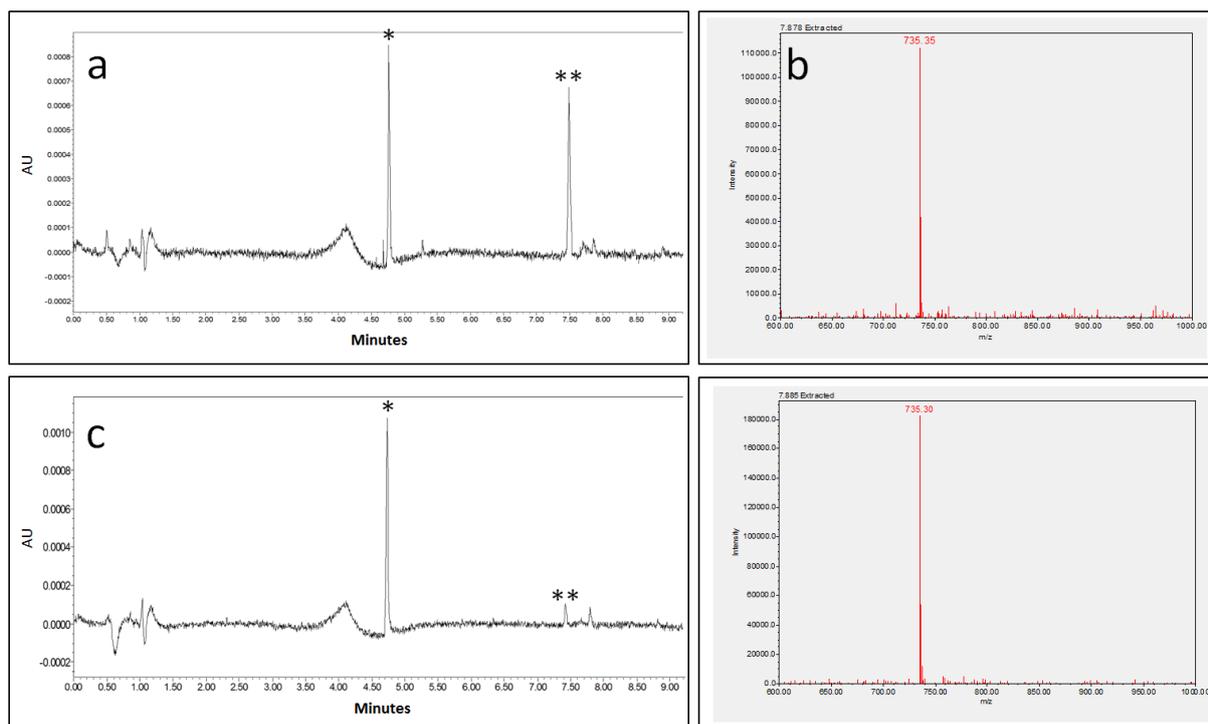


Fig. S24. UPLC-MS of sortase-mediated ligation of tyramine and Abz-LPETGK-Dnp: a) UPLC trace ligation mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 10 mM 4-(Trifluoromethyl)-benzylamine (4-TFB amine, 45% (v/v) DMSO, 20 μ M Sa-SrtA WT); b) Mass spectrum of ligation mixture in (a). The theoretical molecular weight of conjugate (Abz-LPET-4-TFB amine) is 735.32 da. c) UPLC trace ligation mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 4-TFB amine, 45% (v/v) DMSO, 20 μ M Sa-SrtA M3); d) Mass spectrum of ligation mixture in (c). Absorbance was monitored at 355 nm. Peak* is the generated product GK-Dnp. Peak** is the substrate Abz-LPETGK-Dnp. Conversions of Abz-LPETGK-Dnp to Abz-LPETG-4-TFB in (a) and (c) are 54.1 %, 94.3%, respectively. Yield of Abz-LPETGK-Dnp to Abz-LPETG-4-TFB in (a) and (c) are 0.541 and 0.943 mM, respectively.

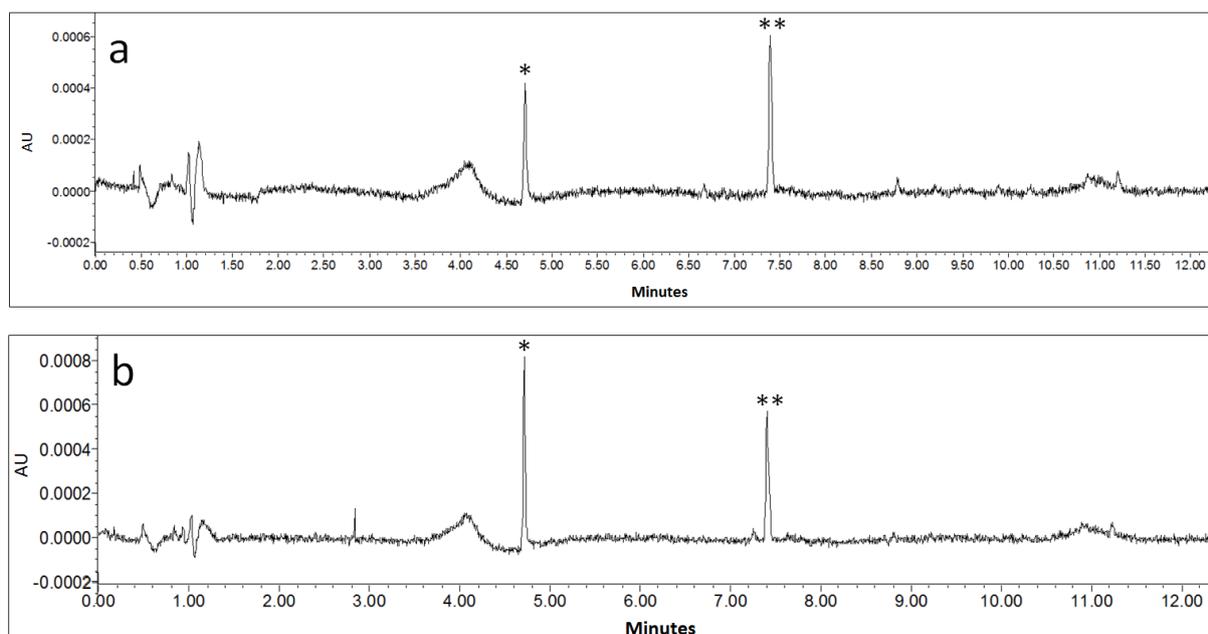


Fig. S25. UPLC of sortase-mediated ligation of O-(2-Aminoethyl) polyethylene glycol (PEG amine, $M_p=5000$ da) and Abz-LPETGK-Dnp: a) UPLC trace of ligation mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 10 mM PEG amine, 45% (v/v) DMSO, 20 μ M Sa-SrtA WT); b) UPLC trace ligation mixture (100 μ L, 1 mM Abz-LPETGK-Dnp, 10 mM PEG amine, 45% (v/v) DMSO, 20 μ M Sa-SrtA M3). Absorbance was monitored at 355 nm. Peak* is the generated product GK-Dnp. Peak** is the substrate Abz-LPETGK-Dnp. Conversions of Abz-LPETGK-Dnp to Abz-LPETG-PEG in (a) and (b) are 36.2 %, 69.6%, respectively. Yield of Abz-LPETGK-Dnp to Abz-LPETG-PEG in (a) and (c) are 0.362 and 0.696 mM, respectively.

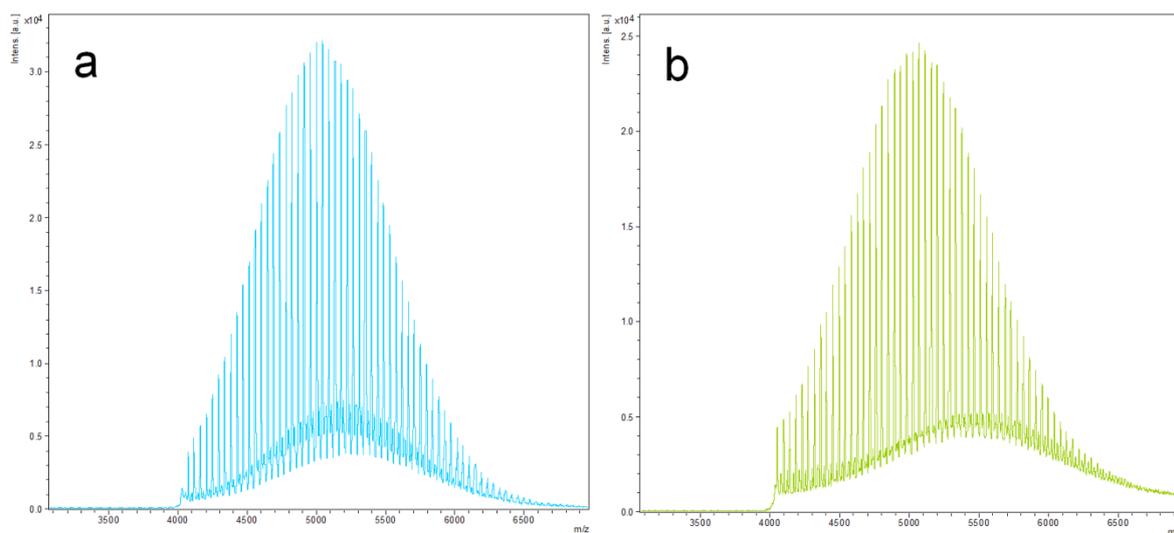


Fig. S26 MALDI-TOF MS of (a) PEG-amine, (b) reaction mixture catalyzed by Sa-SrtA M3 (after 5 h conjugation) in 45% (v/v) DMSO. The average molecular weight of PEG amine is 5000 Da (Dalton) and the average molecular weight of Abz-LPET-PEG is 5545 Da (Dalton). A shift of molecular weight (5100 to 5600 da) of PEG was observed in the bottom part in (b).

References

1. Z. Zou, D. M. Mate, K. Rubsam, F. Jakob and U. Schwaneberg, *Acs Comb Sci*, 2018, **20**, 203-211.
2. F. Cheng, L. L. Zhu and U. Schwaneberg, *Chem Commun*, 2015, **51**, 9760-9772.
3. T. S. Wong, K. L. Tee, B. Hauer and U. Schwaneberg, *Nucleic Acids Research*, 2004, **32**, e26.
4. S. Islam, D. M. Mate, R. Martinez, F. Jakob and U. Schwaneberg, *Biotechnol Bioeng*, 2018, **115**, 1106-1115.
5. A. J. Ruff, T. Kardashliev, A. Dennig and U. Schwaneberg, in *Directed Evolution Library Creation: Methods and Protocols*, eds. E. M. J. Gillam, J. N. Copp and D. Ackerley, Springer New York, New York, NY, 2014, DOI: 10.1007/978-1-4939-1053-3_4, pp. 45-68.
6. M. Blanus, A. Schenk, H. Sadeghi, J. Marienhagen and U. Schwaneberg, *Analytical Biochemistry*, 2010, **406**, 141-146.
7. I. Chen, B. M. Dorr and D. R. Liu, *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 11399-11404.
8. A. Aulabaugh, W. D. Ding, B. Kapoor, K. Tabei, L. Alksne, R. Dushin, T. Zatz, G. Ellestad and X. Y. Huang, *Analytical Biochemistry*, 2007, **360**, 14-22.
9. E. Krieger, G. Koraimann and G. Vriend, *Proteins: Structure, Function, and Bioinformatics*, 2002, **47**, 393-402.
10. J. Van Durme, J. Delgado, F. Stricher, L. Serrano, J. Schymkowitz and F. Rousseau, *Bioinformatics (Oxford, England)*, 2011, **27**, 1711-1712.
11. R. Guerois, J. E. Nielsen and L. Serrano, *Journal of molecular biology*, 2002, **320**, 369-387.
12. Y. Zong, T. W. Bice, H. Ton-That, O. Schneewind and S. V. Narayana, *J Biol Chem*, 2004, **279**, 31383-31389.
13. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845-854.
14. H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Computer Physics Communications*, 1995, **91**, 43-56.
15. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *J Comput Chem*, 2005, **26**, 1701-1718.
16. L. D. Schuler, X. Daura and W. F. Van Gunsteren, *Journal of computational chemistry*, 2001, **22**, 1205-1218.
17. S. Pall, M. J. Abraham, C. Kutzner, B. Hess and E. Lindahl, 2014.
18. C. Oostenbrink, A. Villa, A. E. Mark and W. F. Van Gunsteren, *Journal of computational chemistry*, 2004, **25**, 1656-1676.
19. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Research*, 2004, **32**, W665-W667.
20. T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe and N. A. Baker, *Nucleic Acids Research*, 2007, **35**, W522-W525.
21. A. V. Onufriev and E. Alexov, *Quarterly reviews of biophysics*, 2013, **46**, 181-209.
22. H. Berendsen, J. Grigera and T. Straatsma, *Journal of Physical Chemistry*, 1987, **91**, 6269-6271.
23. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of chemical physics*, 1995, **103**, 8577-8593.
24. J. Norberg and L. Nilsson, *Biophysical Journal*, 2000, **79**, 1537-1553.
25. B. Hess, H. Bekker, H. J. Berendsen and J. G. Fraaije, *Journal of computational chemistry*, 1997, **18**, 1463-1472.
26. V. N. Maiorov and G. M. Crippen, *Proteins: Structure, Function, and Bioinformatics*, 1995, **22**, 273-283.
27. R. Walser, P. H. Hünenberger and W. F. van Gunsteren, *Proteins: Structure, Function, and Bioinformatics*, 2001, **43**, 509-519.
28. M. Y. Lobanov, N. Bogatyreva and O. Galzitskaya, *Molecular Biology*, 2008, **42**, 623-628.
29. A. Luzar, *The Journal of Chemical Physics*, 2000, **113**, 10663-10675.

30. W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577-2637.
31. K. Kulińska, T. Kuliński, A. Lyubartsev, A. Laaksonen and R. W. Adamiak, *Computers & Chemistry*, 2000, **24**, 451-457.
32. F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander and M. Scharf, *Journal of computational chemistry*, 1995, **16**, 273-284.
33. E. Lindahl, B. Hess and D. Van Der Spoel, *Journal of molecular modeling*, 2001, **7**, 306-317.
34. W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33-38.
35. H. Schagger, *Nat Protoc*, 2006, **1**, 16-22.