Supporting Information for:

Molecular Basis of *Streptococcus mutans* Sortase A Inhibition by the Flavonoid Natural Product *trans*-Chalcone

Daynea J. Wallock-Richards[‡,#], Jon Marles-Wright[§,#], David J. Clarke[‡], Amarnath Maitra[Ŧ], Michael Dodds[Ŧ], Bryan Hanley[Ψ] and Dominic Campopiano[‡,*]

[‡]School of Chemistry, The University of Edinburgh, Edinburgh, Scotland, EH9 3JJ, United Kingdom

[§]Institute of Structural and Molecular Biology, The University of Edinburgh, Edinburgh, Scotland, EH9 3JJ, United Kingdom

[Ŧ]Wm. WRIGLEY Jr. Company, 1132 W. Blackhawk Street, Chicago, IL 60642

[Ψ]The Knowledge Transfer Network, The Roslin Institute, Easter Bush, Roslin, EH25 9RG
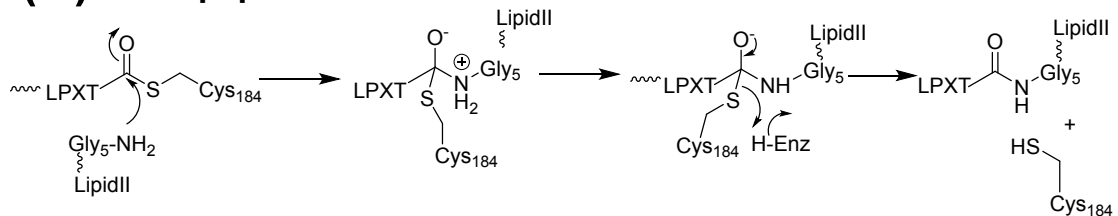
(A) Thioesterification
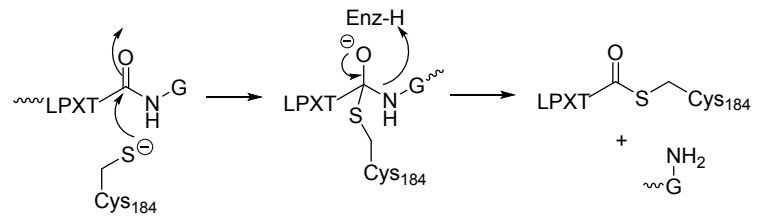
Cleavage of amide bond between threonine and glycine residues of target protein by active site cysteine residue (Cys184)

Oxoanion intermediate which is stabilsed by arginine residue (Arg197).

Formation of a thioacyl intermediate and loss of C-terminus of protein substrate

(B) Transpeptidation

Nucleophilic attack of the thioacyl intermediate by the free amino group of the pentaglycine side chain of lipid II

Formation of a transient tetrehedral intermediate

Collapse of the tetrahedral intermediate

C-terminal linkage of the surface protein to the pentaglycine crossbridge of the lipid II

S1. Proposed reaction mechanism of sortase A. The numbering of residues is based on the *S. aureus* SrtA sequence. The catalytic cycle is a two-step reaction which involves a (A) amide bond cleavage at the conserved LPXTG motif and thioesterification of Cys184 followed by (B) transpeptidation and amide bond formation between the protein target and lipid II.

```
SmSrtA                                               TT   TT            TT
        1       10        20        30        40              50        60
SmSrtA  MKKERQSRKKRSFLRTFLPILLLVI..GLALIFNTPIRNTLIA....WNTNRY..QVSNVSKKDIEHNK
SpSrtA  ......................................................................
SaSrtA  ........MKKWTNRLMTIAGVVLILVAAYLF....AKPHID..............NYLHDKDKD...
SaSrtB  ..............................................................MEDKQER
BaSrtB  ........................SNAIFM.......................DYYENRKVM
SpnSrtC ....................................ES.NQ.QI....ADFDKEKATLDEADIDERM
SagSrtC ...........................GSFTES.NN.QT....QDFERAAKKLSQKEINRRM
SsSrtC  ..........................................................MRI
AoSrtC  ...............MPMSDYDIPTTENLYFQGAMATQHNNDEQQRLAKMYTATVNSAGPETIAKER

SmSrtA       TT   0000000000.........00.0000000000000   β1   η1   β2        η2
               70        80        90      100       110
SmSrtA  AAHSSFDFKKVESIS..........TQ.SVLAAQMAAQKLPVIGGIAIPDLKINLPIFKGLDNV.GLTY
SpSrtA  ..........MLVP..........RG.SVLQAQMAAQQLPVIGGIAIPELGINLPIFKGLGNT.ELIY
SaSrtA  EKIEQYDKNVKEQASKD....KKQ.....QAKPQIPKDKSKVAGYIEIPDADIKEPVYPGPATPEQLNR
SaSrtB  ANYEKL.......QQKFQMLMSKHQAHVRPQFESLEKINKDIVGWIKLSGTSLNYPVLQGKTNHDYLNL
BaSrtB  AEAQNI.......YEKSPMEEQSQDGEVRKQFKALQQINQEIVGWITMDDTQINYPIVQAKDNDYYLFR
SpnSrtC KLAQAFNDSLNNVVSGDPWSEEMKKKG..RAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEE.VLQQ
SagSrtC ALAQAYNDSLNNVHLEDPYEKKRIQKG..IAEYARMLEVSEKIGIISVPKIGQKLPIFAGSSQE.VLSK
SsSrtC  DLAQAFNESLNNVVSEDPYTKSRHEAG..RVEYARMLELHEKIGYVEIPKIEVKLPVYAGTSET.VLQK
AoSrtC  ASAETYNNNLESAPILDPWLESQRPDTPQYQAYLHEMDIDPVMARIVIPSIHVSLPIYHGTDSR.TLTE

SmSrtA    β3     TT        TT     β4      η3      η4      TT      β5
        120       130       140       150       160            170
SmSrtA  G..AGTMKNDQV.......MG.EN.NYALASHHVFGMTGSSQMLFSPLERAKEGMEI......YLTDKN
SpSrtA  G..AGTMKEEQV.......MGGEN.NYSLASHHIFGITGSSQMLFSPLERAQNGMSI......YLTDKE
SaSrtA  G..VSFAEENES.......LDDQNISIAGHTFI..DRPNYQFTNLKAAKKGSMV......YFKVGN
SaSrtB  DFEREHRRKGSIFMDFRNELKNLNHNTILYGHHVG....DNTMFDVLEDYLKQSFYEKHKIIEFDNKY
BaSrtB  NYKGEDMRAGSIFMDYRNDVKSQNRNTILYGHRMK.....DGSMFGSLKKMLDEEFFMSHRKLYYDTLF
SpnSrtC G..AGHLEGTSLP......IGGNSTHAVITAHTGL...P.TAKMFTDLTKLKVGDKF......YVHNIK
SagSrtC G..AGHLEGTSLP......IGGNSTHTVITAHSGI...P.DKELFSNLKKLKKGDKF......YIQNIK
SsSrtC  G..VGHLEGTSLP......IGGSDTHTVLTAHTGL...P.KARLFTDLTKVKIGDTF......YIHNIV
AoSrtC  G..VGHLFGTSLP......VGGPSTHSVLTGHTGL...S.TATMFDNLNQLKKGDVF......YVSSLG
                                        ▲

SmSrtA      β6          TT              η5        TT     β7         TT
          180                      190       200       210
SmSrtA  KVYTYVISEVKTVTPE......H..................VEVIDNRPGQN.EVTLVTCTDAG.AT.
SpSrtA  KIYEYIIKDVFTVAPE......R..................VDVIDDTAGLK.EVTLVTCTDIE.AT.
SaSrtA  ETRKYKMTSIRDVKPT......D..................VGVLDEQKGKDKQLTLITCDDYN.EKT
SaSrtB  GKYQLQVFSAYKTTTKDNYIRTDFENDQDYQQFLDETKRKSVINSDVNVTVKDKIMTLSTCEDAYSETT
BaSrtB  EGYDLEVFSVYTTTTDFYYIETDFSSDTEYTSFLEKIQEKSLYKTDTTVTAGDQIVTLSTCDYALDPEA
SpnSrtC EVMAYQVDQVKVIEPT......N..................FDDLLIVPGHD.YVTLLTCTPYM.INT
SagSrtC ETIAYQVDQIKVVTPD......N..................FSDLLVVPGHD.YATLLTCTPIM.VNT
SsSrtC  ETLAYEVDQIVVAEPT......Q..................FEELLVKPGQD.YATLLTCTPYM.VNT
AoSrtC  QTLKYEVNDITVVKPE......E..................TDSLRKVPGRD.LVTLITCTPYG.VNS
                                                              ▲

SmSrtA      β8      η6      α2
          220       230       240
SmSrtA  ARTIVHGTYKGETDFNKTSKKIKKAFRQSYNQISF
SpSrtA  ERIIVKGELKTEYDFDKAPADVLKAFNHSYNQVST
SaSrtA  GRWEKRKIFVA.TEVK...............
SaSrtB  KRIVVVAKIIKVS.................
BaSrtB  GRLVVHAKLVKRQ.................
SpnSrtC HRLLVRGHRIP.YVAEVEEEFIA.A.....NKLSH
SagSrtC HRLLVRGHRIP.YKGPIDEKLIK.D.....GHLNT
SsSrtC  HRLLVRGHRIP.YVAEEHLE..............
AoSrtC  HRLLVTGERVP.MDPT...............
         ▲
```

S2. Multiple amino acid sequence alignment of sortase ezymes (SrtA, SrtB and SrtC) with structures determined. Identical residues are highlighted in red, similar residues are coloured in red and highlighted with a blue box. Catalytic triad identified in srtA are indicated with a blue triangle. The PDB code for the sequences of the sortase enzymes are as follows: *Streptococcus mutans* sortase A (SmSrtA) 4TQX; *Streptococcus pyogenes* sortase A (SpSrtA), 3FN5; *Staphylococcus aureus* sortase A (SaSrtA), 1T2P; *Staphylococcus aureus* sortase B (SaSrtB), 1NG5; *Bacillus anthracis* sortase B (BaSrtB), 2OQW; *Streptococcus pnemoniae* sortase C-1 (SpnSrtC), 2W1J; *Streptococcus agalactiae* sortase C-1 (SagSrtC), 3O0P; *Streptococcus suis* sortase C-1 (SsSrtC), 3RE9; *Actinomyces oris* sortase C-1 (AoSrtC), 2XWG. The alignment was ran by Clustal Omega of EMBL-EBI and ESPript 3.x.

S3. Gene Cloning and Mutagenesis. The gene sequence for *S. mutans* sortase A (SmSrtA WT) was downloaded from the GeneBank database (accession number AB089932) and used as a template for designing a synthetic gene (GenScript, cloned in pUC57), with the DNA sequence modified to remove the first membrane targeting sequence of 120 bp/40 amino acids (we refer to internally as SmSrtAΔ_{N40}) and insert the restriction endonuclease sites Nco1(5') and Xho1(3').

This allowed for the cloning of the synthetic gene into a pET28a expression plasmid. Overexpression of pET28a/SmSrtA produced an untagged recombinant SrtA protein of 206 amino acids with an N-terminal truncation of the first 40 amino acids. This clone was then used to prepare the SrtA H139A mutant using the Liu mutagenesis protocol[1] with the following primers:

5'-gctagc**gct**catgttttttggtatgaccg-3' (Forward)
5'-aaaacatg**agc**gctagcaagagcataa-3' (Reverse)

The bases mutated are shown in bold. DNA sequencing was used to verify the SrtA H139A mutant. Mass spectrometry analysis showed that both wild-type and mutant SrtA were expressed with the N-terminal methionine removed.

S4. Expression and Purification of SrtA and H139A Mutant. A single colony of pET28a/SmSrtA BL21 (DE3) (or pET28a/SmSrtA H139A mutant) was picked into 500 mL of LB broth supplemented with kanamycin (30 µg/mL) and grown overnight at 37 ^{O}C. The overnight culture was used to inoculate fresh LB and the cells were grown to an OD_{600} of 0.6 before induction with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. The cultures were incubated for a further 5 hours at 20 ^{o}C, 200 rpm. The cells were harvested by centrifugation at 5000 ×g for 15 minutes. The cell pellet was resuspended in buffer A (20 mM MES.NaOH, pH 6.0, 20 % glycerol) and lysed by sonication (Soniprep 150, 15 min cycles with 30 s intervals). The resulting crude lysate was centrifuged at 60076 × g for 30 min at 4 ^{O}C. The cell free extract was loaded onto MonoS 5/50 cation exchange column (GE Healthcare) and eluted with increasing concentration of buffer B (20 mm MES.NaOH, pH6.0, 20% glycerol, 2M NaCl). Fractions containing the protein of interest were pooled and further purified by size exclusion chromatography using a Superdex S200 column 16/60 (GE Healthcare) equilibrated with 20 mM MES.NaOH, pH 6.0, 150 mM NaCl. Further oligomeric analysis was carried out by SDS PAGE and mass spectrometry.
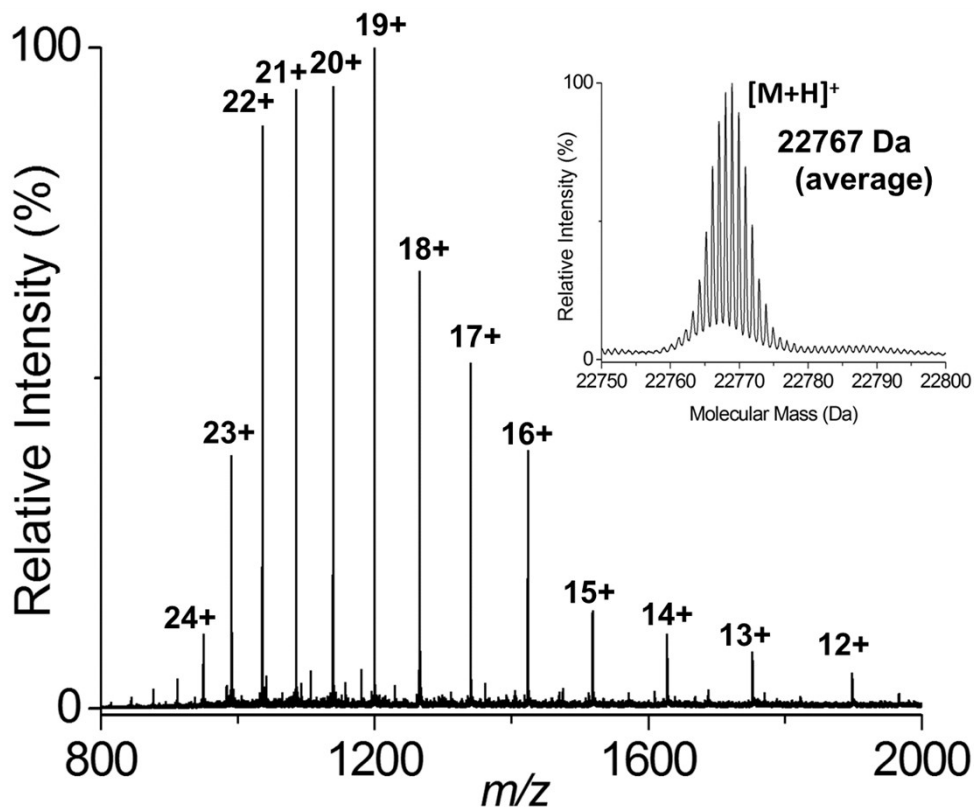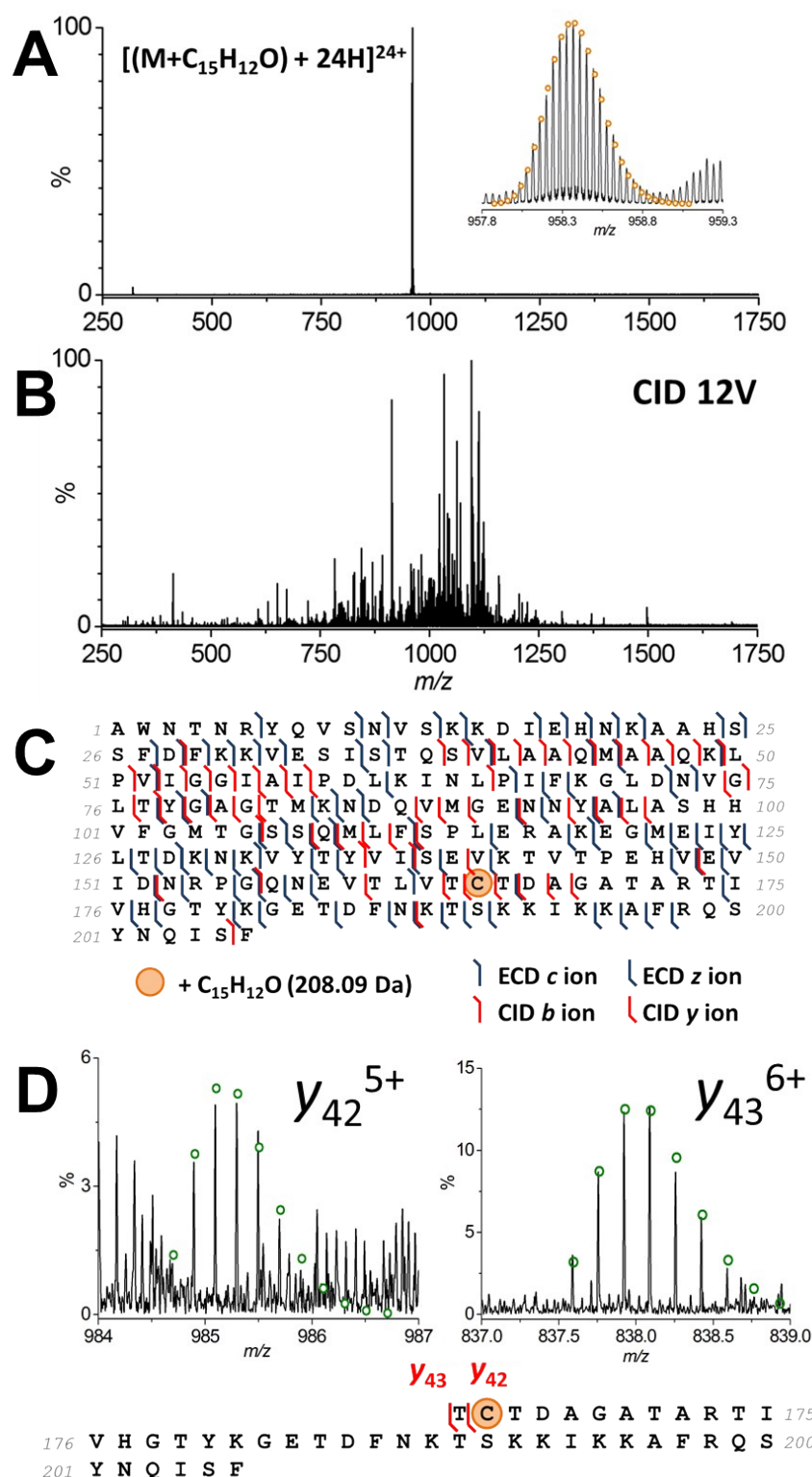
S5. Mass spectrum of recombinant *S. mutans* Sortase A (SrtA). LC-MS experiments were performed using a Synapt G2 (Waters) instrument equipped with an Acquity UPLC and an electrospray source. Separation was achieved using a reverse phase C4 Aeris Widepore 50 × 2.1 mm HPLC column (Phenomenex) and a gradient of 5-95% MeOH (0.1% Formic Acid) over 10 minutes. The main figure shows the charge state distribution obtained for SmSrtA. This corresponds to a deconvoluted molecular mass of 22767 Da, ([M+H]$^+$) see insert, which was consistent with the predicted theoretical mass based on the amino acid sequence of the protein (theoretical [M+H]$^+$, 22767 Da).
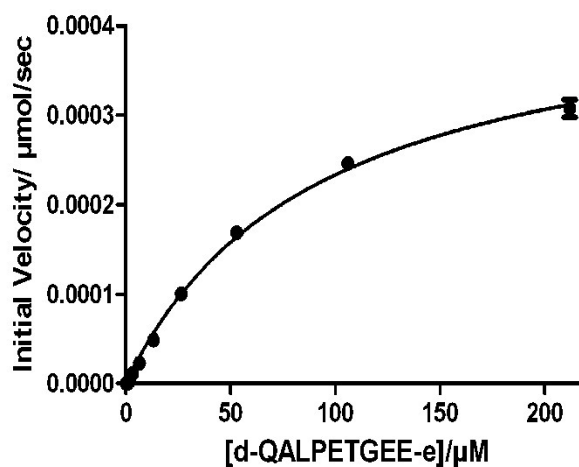
Enzyme	$K_m/\mu M$	$V_{max}/\mu Ms^{-1}$
SmSrtA	90.4 ± 4.7	$4.46 \pm 0.11 \times 10^{-4}$

S6. Determination of K_m and V_{max} for *S. mutans* SrtA. SrtA activity was measured using a previously published method that uses a fluorometric assay to monitor the hydrolysis of a fluorogenic peptide.[2] The peptide substrate 4-(4-dimethylaminophenylazo)benzoicacid(Dabcyl)-QALPETGEE-5-[(2-aminoethyl) amino]naphthalene-1-sulfonic acid (Edans) (Dabcyl-QALPETGEE-Edans) was dissolved in dimethyl sulfoxide (DMSO) to prepare a stock concentration of 3.39 mM. This was then diluted with sterile distilled water so that the final concentration of DMSO in the reaction was kept below 5%. The assay was carried out at 37 °C (in triplicate) in reaction buffer (20 mM MES.NaOH, 125 mM NaCl, 2 mM DTT and 5 mM triglycine, pH 6.5). Wells contained 5 µM purified enzyme and varying concentrations of peptide substrate (0.8 -212 µM) in a final volume of 200 µl. Assays were performed in a 96-well plate and peptide cleavage was monitored as an increase of fluorescence using a Biotek Synergy HT plate reader. The increase in fluorescence intensity was recorded as a function of time using an excitation wavelength at 360 nm and recording emission at 485 nm. Changes in fluorescence were converted to molar velocities using a calibration curve of an equimolar mixture of free dabcyl and edans which had a gain of 23 F.U./µmol upon conversion of substrate to product. The initial velocities for substrate cleavage were determined from the progress curves at various substrate concentrations and fitted to the Michaelis-Menten equation to determine the kinetic parameters K_m and V_{max} of the first step (peptide cleavage) using GraphPad Prism 6 software. All reactions were performed in triplicates and the mean values reported.

S7. Inhibition studies of S. mutans SrtA with *trans*-chalcone. The inhibitory activity of SrtA by *trans*-chalcone (IUPAC Name: (E)-1,3-diphenylprop-2-en-1-one) was determined by measuring the rate at which the synthetic fluorescent peptide substrate Dabcyl-QALPETGEE-Edans was cleaved using the fluorescent assay described previously.[3] The reactions were performed on a 200 µL scale containing reaction buffer (20 mM MES.NaOH, pH 6.5 and 125 mM NaCl), triglycine (5 mM), peptide substrate (25 µM), 5 µM recombinant SrtA and varying concentrations of *trans*-chalcone (1-100 µM). *trans*-chalcone was dissolved in DMSO and added to reaction mixture so that the final concentration of DSMO was >1%. Appropriate blanks containing all of the above reagents except for *trans*-chalcone were used as a negative control and morin and curcumin (known natural product sortase A inhibitors) were used as a positive controls. All reactions were done in triplicates. The inhibitory activity was defined as the concentration that resulted in 50% inhibition of enzyme activity relative to 1% DMSO.

A $[(M+C_{15}H_{12}O) + 24H]^{24+}$

957.8 958.3 *m/z* 958.8 959.3

B CID 12V

C

```
  1 A W N T N R Y Q V S N V S K K D I E H N K A A H S 25
 26 S F D F K V E S I S T Q S V L A A Q M A A Q K L 50
 51 P V I G G I A I P D L K I N L P I F K G L D N V G 75
 76 L T Y G A G T M K N D Q V M G E N N Y A L A S H H 100
101 V F G M T G S S Q M L F S P L E R A K E G M E I Y 125
126 L T D K N K V Y T Y V I S E V K T V T P E H V E V 150
151 I D N R P G Q N E V T L V T C T D A G A T A R T I 175
176 V H G T Y K G E T D F N K T S K K I K K A F R Q S 200
201 Y N Q I S F
```

⬤ + C$_{15}$H$_{12}$O (208.09 Da)

⌐ ECD *c* ion   ⌐ ECD *z* ion
⌐ CID *b* ion   ⌐ CID *y* ion

D

y_{42}^{5+}

984 985 986 987 *m/z*

y_{43}^{6+}

837.0 837.5 838.0 838.5 839.0 *m/z*

y$_{43}$ *y*$_{42}$

```
         T C T D A G A T A R T I 175
176 V H G T Y K G E T D F N K T S K K I K K A F R Q S 200
201 Y N Q I S F
```

S8. Mapping the location of the chalcone modification in SrtA by top down mass spectrometry.
(A)   Typical mass spectrum of a single isolated charge state of chalcone-modified SrtA. Here the 24+ charge state is shown. *Inset,* Isotope distribution of the 24+ species. The ion has an isotope pattern consistent with Michael addition of chalcone (+C$_{15}$H$_{12}$O). The simulated isotope pattern is displayed in orange ([C$_{1023}$H$_{1634}$N$_{276}$O$_{311}$S$_7$]$^{24+}$).
**(B)**   Typical top-down mass spectrum. Here the CID spectrum of the 24+ species is shown.

8

(C) Combined top down fragment map showing all of the assigned product ions from all fragmentation experiments. Fragments resulting from CID are displayed in red, fragments resulting from ECD are displayed in blue. Fragmentation allowed the assignment of *S. mutans* SrtA Cys205 as the residue modified by chalcone.

(D) The isotopic distributions of the y_{42}^{5+} ($[C_{207}H_{333}N_{60}O_{64}S_1(C_{15}H_{12}O)]^{5+}$) and y_{43}^{6+} ($[C_{211}H_{341}N_{61}O_{66}S_1(C_{15}H_{12}O)]^{6+}$) fragment ions which allow the assignment the site of modification as Cys205. The simulated isotope patterns are shown as scatterplots in green.
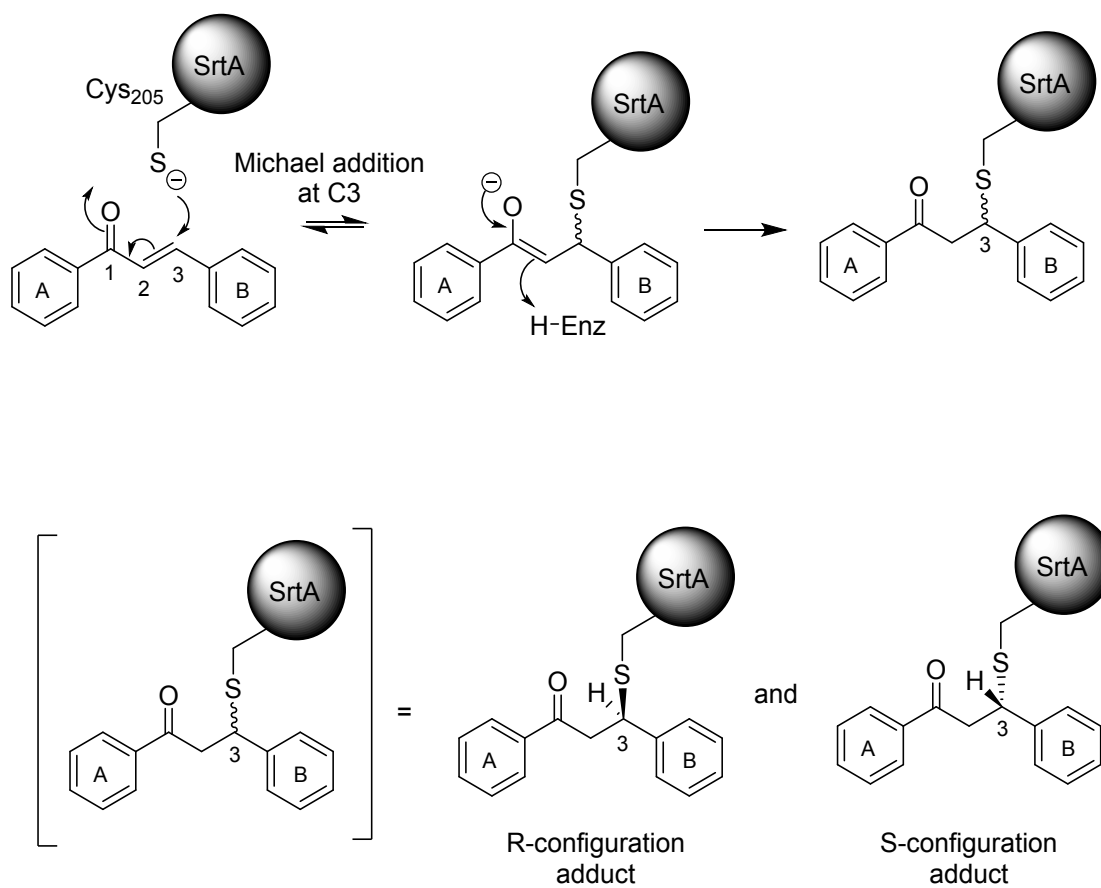
S9. Proposed mechanism for the formation of the SrtA-*trans*-chalcone adduct.
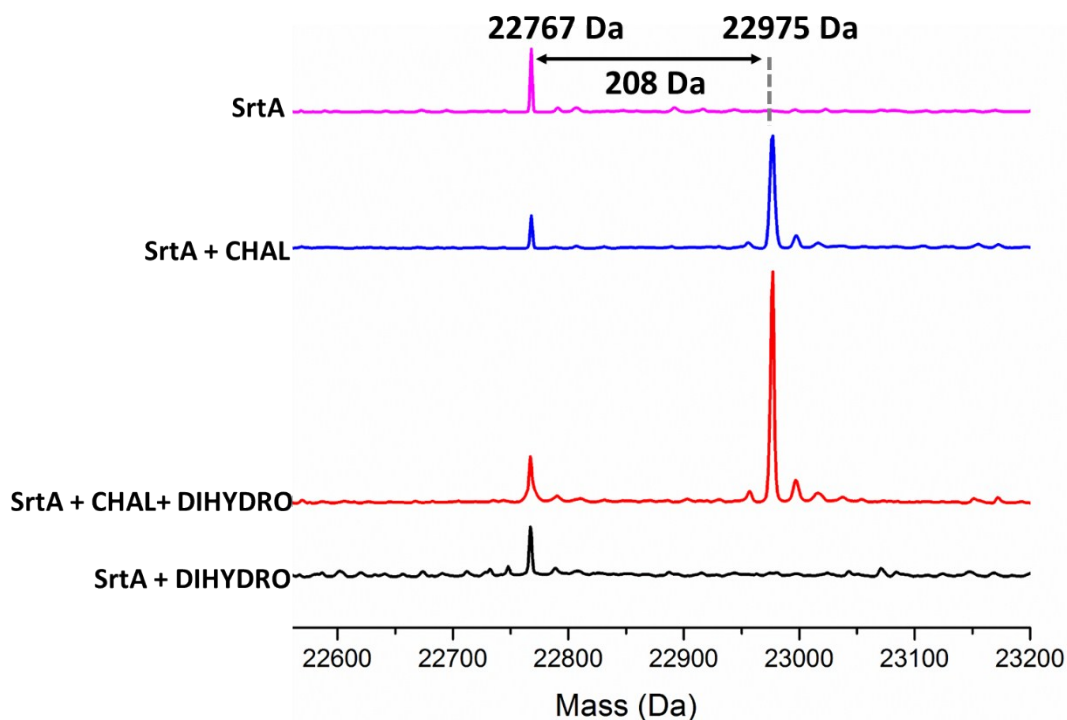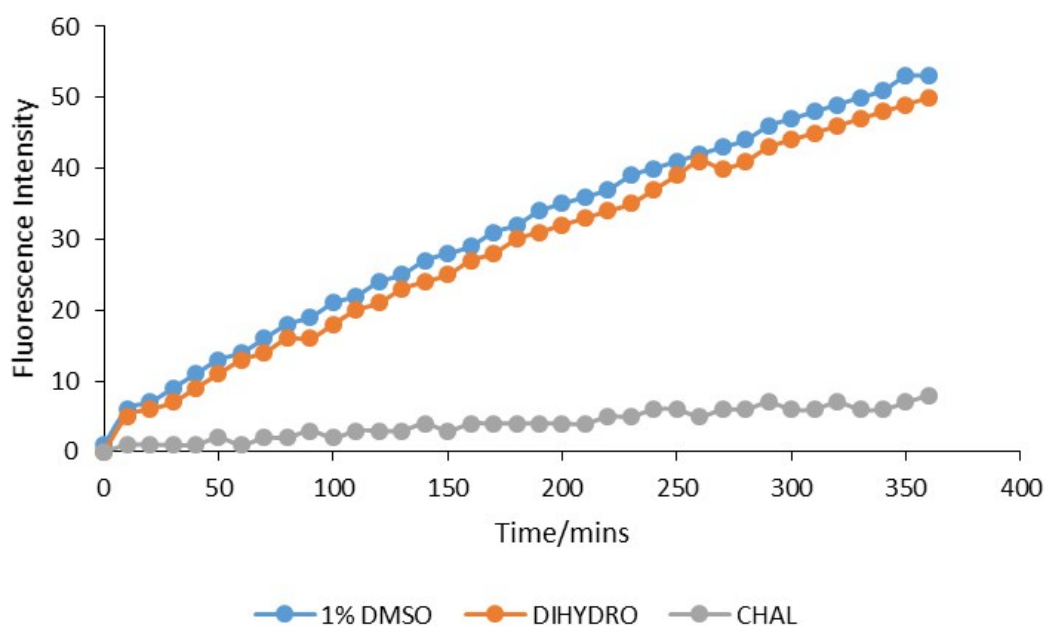
S10. LC ESI-MS analysis and inhibition of SrtA incubated with *trans*-chalcone and dihydrochalcone. SrtA (5 µM) was incubated with (i) *trans*-chalcone (100 µM), (ii) dihydrochalcone (100 µM) and (iii) mixture of *trans*-chalcone (100 µM) and dihydrochalcone (100 µM) for 16h.
**(A)** Mass spectrometry analysis of reactions (i) to (iii) by LC-MS. A SrtA adduct was observed with trans-chalcone but not the dihydrochalcone.
(B) Enzyme assay data obtained from SrtA incubated with *trans*-chalcone (100 µM) and dihydrochalcone (100 µM) for 16h and analysed by FRET assay. The dihydrochalcone displayed no detectable activity.

S11. Effect of *trans*-chalcone on biofilm formation by *S. mutans*. *S. mutans* biofilm was grown using the ACTA Active Attachment (AAA) model.[4] In this model 24 glass slides are assembled on to a steel lid by plastic pegs. The steel lid assembly can fits into a standard Corning 24-well plate. The glass slides are used for biofilm attachment of *S. mutans* cells and the subsequent growth of biofilm.

S. mutans culture was grown overnight in BHI (10 mL) with *trans*-chalcone (0-500 µM) at 37 °C. The AAA apparatus *i.e.* glass slides suspended from a steel lid was incubated with sterile 50-fold diluted saliva (1.5 mL) in a 24-well Corning plate for 2h at 37 °C. The glass slides were then incubated in BHI (1.5 mL) containing 0.2% sucrose, *trans*-chalcone (0-500 µM), and the overnight *S. mutans* inoculum diluted to a final OD_{600} for 2.5 h with moderate shaking (100 rpm) using a 24-well plate format. After inoculation of the glass slides, they were placed in fresh BHI media (supplemented with 0.2 %sucrose and *trans*-chalcone) for biofilm growth for a period of 16-18h.

After 18h of biofilm growth, the glass slides with attached *S. mutans* biofilm were stained with 0.05% crystal violet solution (1.5 mL) for 10 mins. Excess stain was removed by washing with phosphate-buffered saline (PBS). The glass slides were allowed to air dry and then destained in 95% ethanol (2 mL). The absorbance of each well solution (which contained crystal violet stain dissolved in 95% EtOH) was then measured at 600 nm. The absorbance of the well solution is assumed to be directly proportional to biomass/biofilm formation on each glass slide. See below.
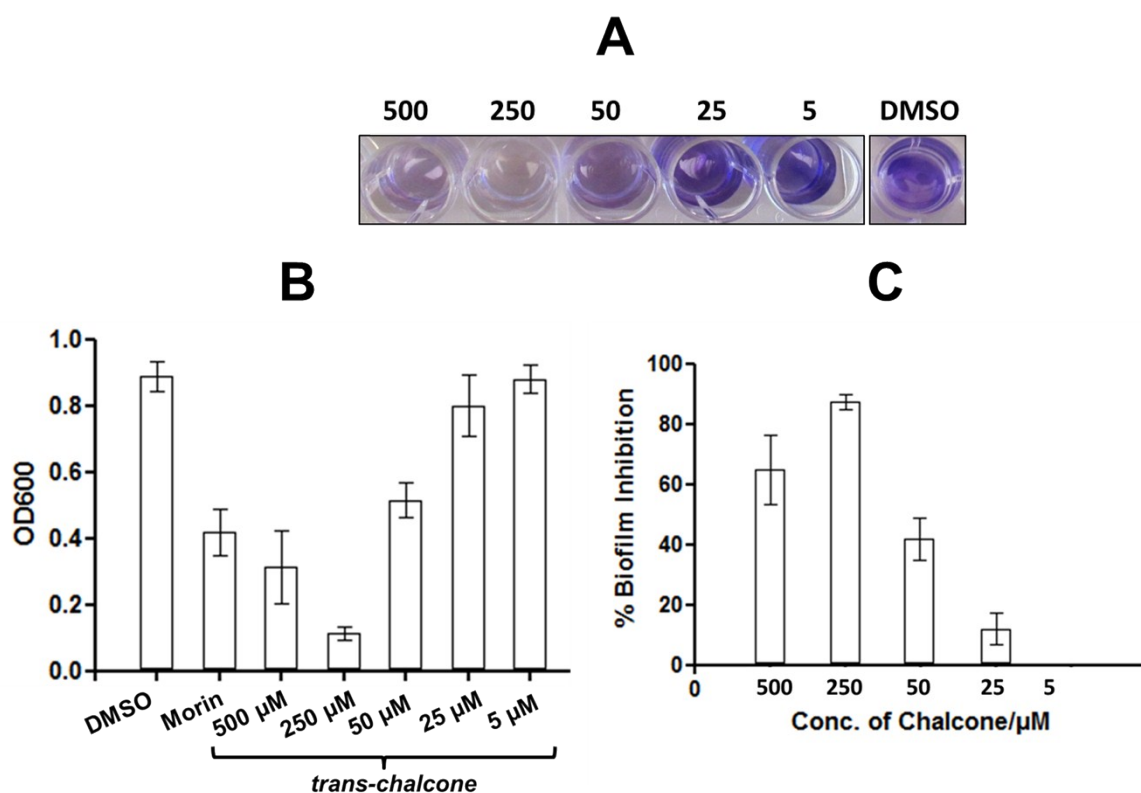

(A) Photograph of crystal violet stain dissolved in EtOH after destaining of *S. mutans* biofilms grown on saliva-coated glass slides in the presence of *trans*-chalcone using the AAA model. Cells were grown in BHI media containing *trans*-chalcone (5-500 µM). A stronger purple colour indicates that more biomass was present/stuck on the glass slide before destaining as is seen in the wells containing the DMSO control. 25 µM and 5 µM *trans*-chalcone.
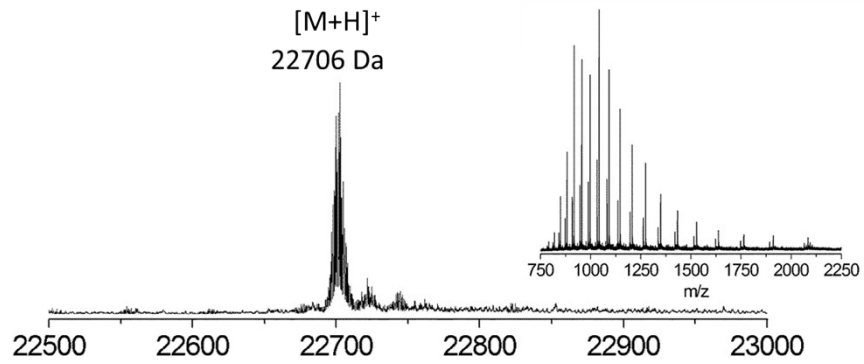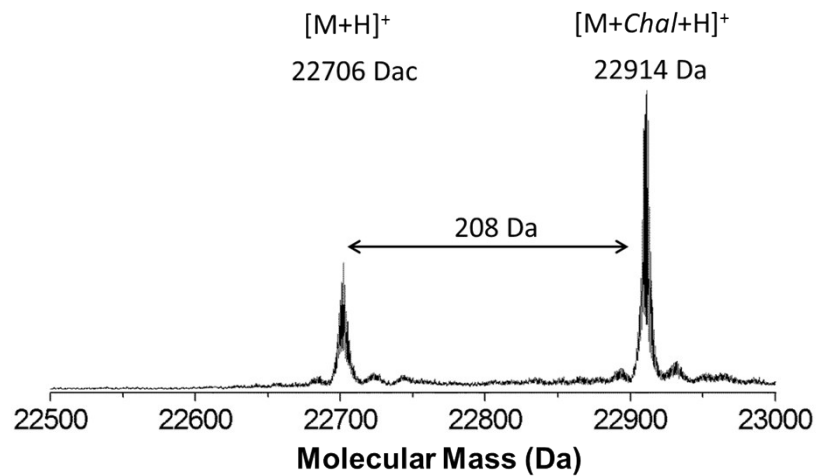**(B)** Quantification of biofilm mass. The optical density of crystal violet stain dissolved in EtOH after destaining was read at 600 nm. The values represent the average experiments performed in triplicate.

12

DMSO was used as the negative control and the known *S. mutans* biofilm inhibitor morin (50 µM) as the positive control.

(C) Inhibition of biofilm formation in the presence of *trans*-chalcone. Biofilm formation was reduced in the presence of *trans*-chalcone in a concentration dependent manner up to 250 µM, with efficacy tailing off at higher concentrations. % values are relative to the DMSO control.
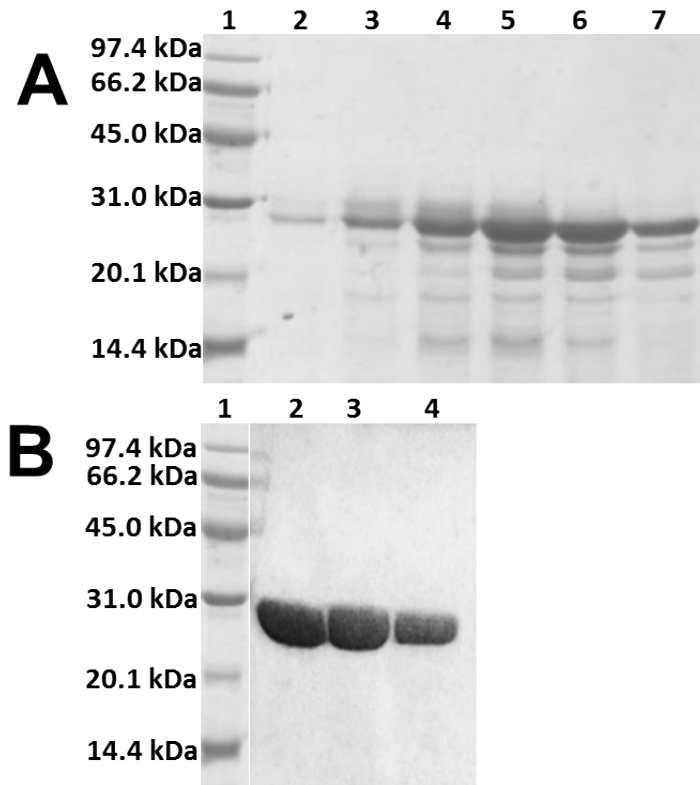
S12. SDS-PAGE analysis of recombinant *S. mutans* SrtA and SrtA H139A after purification by size exclusion chromatography. **(A)** Wild-type SrtA is unstable and undergoes auto proteolysis upon purification. Lane 1: low molecular weight marker; lanes 2-7: Fractions of purified protein from gel filtration on Superdex S200 16/60 (GE Healthcare) 20 mM MES.NaOH, 250 mM NaCl, pH 6.5. **(B)** The SrtA H139A mutant is more stable. Lane 1: low molecular weight marker, lanes 2-4: Fractions of purified protein from gel filtration on Superdex S200 16/60 (GE Healthcare) in 20 mM MES.NaOH, 250 mM NaCl, pH 6.

A

[M+H]+
22706 Da

750 1000 1250 1500 1750 2000 2250
m/z

22500   22600   22700   22800   22900   23000

B

[M+H]+
22706 Dac

[M+*Chal*+H]+
22914 Da

← 208 Da →

22500   22600   22700   22800   22900   23000

Molecular Mass (Da)

S13. Mass spectra of SrtA H139A mutant before and after incubation with *trans*-chalcone.
(A) Purified SrtA H139A (5 μM) was analysed by LC ESI-MS and found to give a mass which corresponds to the predicted monomeric mass of 22705 Da.
(B) SrtA H139A (5 μM) was incubated overnight with *trans*-chalcone (100 μM), followed my LC-MS analysis. A peak at 22914 Da was observed which corresponds to the enzyme-chalcone adduct. MS-MS analysis confirmed the site of modification was Cys205 (data not shown).

S14. Protein crystallisation and data collection, structure solution and analysis. The stable SrtA H139A mutant enzyme was used for structural studies. Purified SrtA H139A in 250 mM NaCl, 20 mM MES.NaOH pH 6.5 was concentrated to 11 mg/ml using a 10 kDa MWCO centrifugal filter device (Sartorius). Concentrated protein was crystallised by hanging drop vapour diffusion in drops of 2 μl protein plus 2 μl crystallisation solution (30 % w/v PEG 4000, NaAc pH 4.6, 0.2 M $(NH_4)_2SO_4$), over 1 ml of the latter. Crystals were obtained in 5 days; these were harvested from the well using a LithoLoop (Molecular Dimensions Limited), transferred briefly to a cryoprotection solution containing well solution supplemented with 20 % w/v PEG200, and the crystals were subsequently flash cooled in liquid nitrogen. All crystallographic datasets were collected on beamlines I24 and I04 at Diamond Light Source (Didcot, UK) at 100 K using Pilatus 6M detectors. Diffraction data were integrated and scaled using XDS and symmetry related reflections were merged with Aimless.[5] Data collection statistics are shown in Table 1. The resolution cut off used for structure determination and refinement was determined based on the $CC_{1/2}$ criterion proposed by Karplus and Diederichs.[6]

The structure of SrtA H139A was determined by molecular replacement using PDBID: 3FN5[7] as the search model, this was modified to match the sequence of the target protein using Chainsaw. A single solution comprising a monomer in the asymmetric unit was found using Phaser.[8] The initial model was rebuilt using Phenix.autobuild[9] followed by cycles of refinement with Phenix.refine[9] and manual rebuilding in coot.[10] The final model was refined with anisotropic B-factors for the protein chain and isotropic B-factors for ligands and water molecules. The model was validated using MolProbity.[11] Structural superimpositions were calculated using Coot.[12] Crystallographic figures were generated with PyMOL.[13] Refinement statistics are shown in Table 1.

	SrtA H139A
Wavelength (Å)	0.92
Resolution range (Å)	47.44 - 1.37 (1.419 - 1.37)
Space group	P 42 21 2
Unit cell	82.822 82.822 57.87 90 90 90
Total reflections	1094198 (109542)
Unique reflections	42782 (4216)
Multiplicity	25.6 (26.0)
Completeness (%)	99.85 (99.98)
Mean I/sigma(I)	25.19 (1.12)
Wilson B-factor	19.25
R-merge	0.08007 (2.888)
R-meas	0.08172
CC1/2	1 (0.618)
CC*	1 (0.874)
Reflections used for R-free	5%
R-work	0.1450 (0.2806)
R-free	0.1792 (0.3296)
Number of non-hydrogen atoms	1837
macromolecules	1585
ligands	22
water	230
Protein residues	201
RMS(bonds)	0.009
RMS(angles)	1.22
Ramachandran favored (%)	98
Ramachandran allowed (%)	2
Ramachandran outliers (%)	0
Clashscore	1.24
Average B-factor	27.50
macromolecules	25.80
ligands	49.60
solvent	37.30

Statistics for the highest-resolution shell are shown in parentheses.

S14. Data Collection and Refinement Statistics.
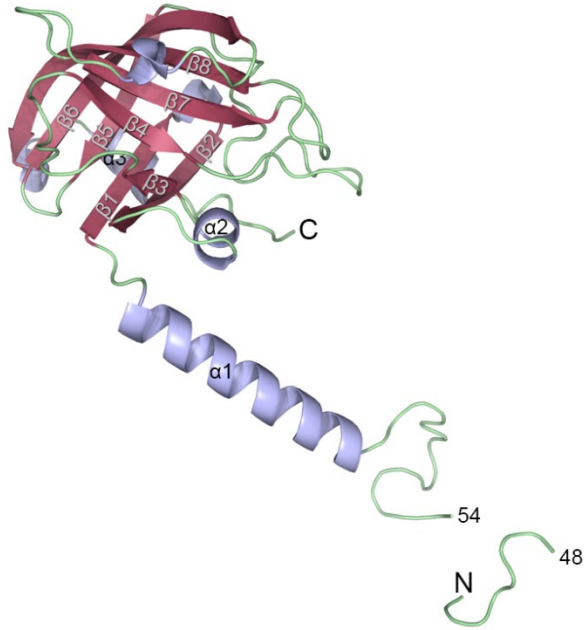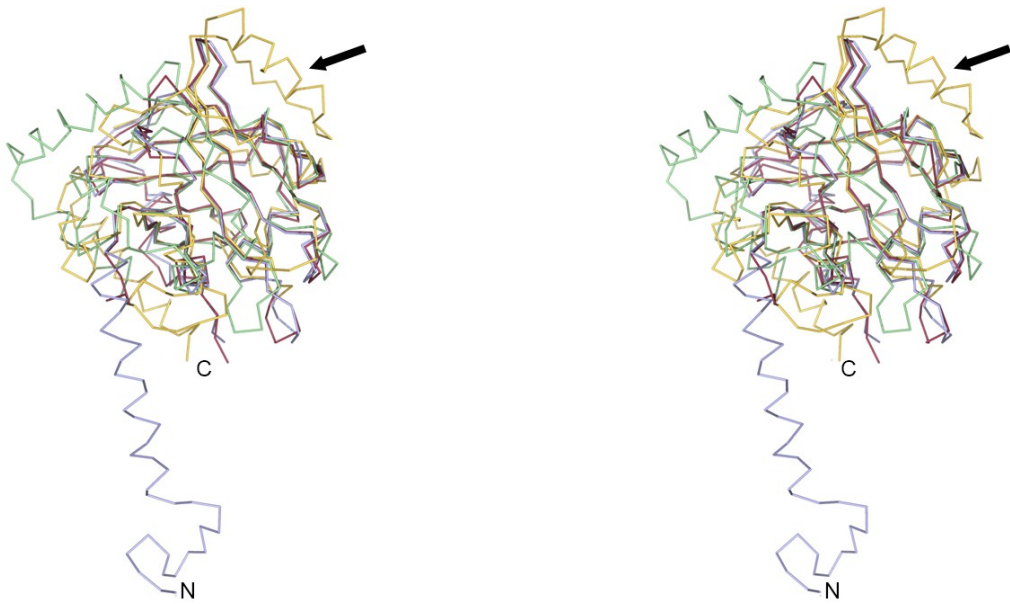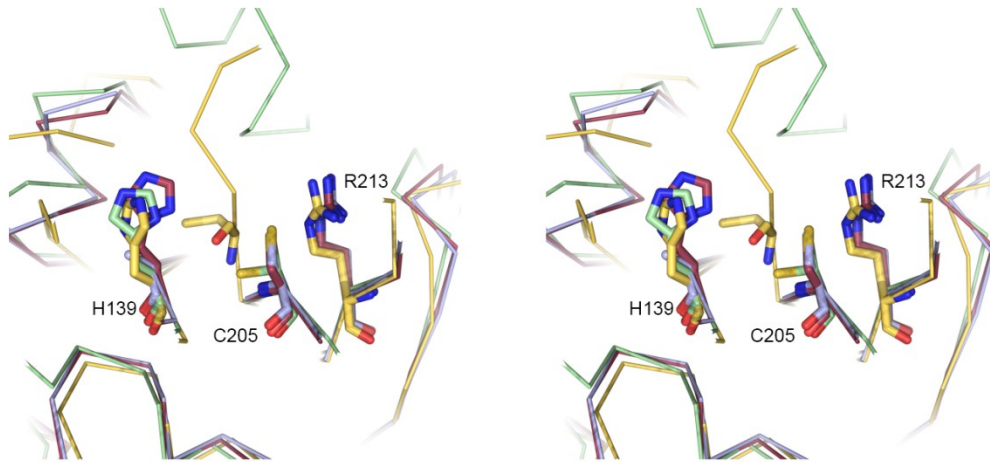
S15. Overlay of the crystal structures of sortase enzymes.
(**A**) Structure of SmSrtA H139A showing the secondary structural elements labelled from the N- to C-terminus and highlighting the extended N-terminal domain (α1) that is unique to this structure.
(**B**) Stereo view of an overlay of four sortase structures showing the highly conserved core fold. (*Sm*SrtA, 4TQX, blue; *Sp*SrtA, 3FN5, raspberry; *Sa*SrtB, 1NG5, yellow; *Spn*SrtC, 2W1J, green). The *Sm*SrtA reported in our work has a unique N-terminal helix arrangement which is distinct to those seen for the other sortase family enzymes, N-terminus of *Sm*SrtA indicated with an 'N' and *Sa*SrtB indicated with an arrow.
(**C**) Stereo view of the active site residues of the four sortase structures as in (**B**), with residue numbers labelled for the *Sm*SrtA protein. The only significant difference in active residues is seen between the *Sa*SrtB active site cysteine and the other sortase family proteins, where it is shifted by approximately 4 Å due to a loop movement (in yellow) in the active site.

S16. Modelling of the SrtA *trans*-chalcone Michael adduct. Both the R- and S- form of the trans-chalcone cysteine Michael adduct were built and their geometry was optimized using Jligand.[14] Geometry description files were output as cif format files and Cys205 was modified in coot with both forms of the trans-chalcone. The His139 residue was modeled to minimize steric clashes with other side-chains and the ligand. The resulting models were minimized using refmac and figures produced using Pymol.[13]

Supplementary references

(1)   Liu, H.; Naismith, J. H. *BMC Biotechnol* **2008**, *8*, 91.

(2)   Ton-That, H.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. *J Biol Chem* **2000**, *275*, 9876.

(3)   Hu, P.; Huang, P.; Chen, W. M. *Appl Biochem Biotechnol* **2013**, *171*, 396.

(4)   Exterkate, R. A.; Crielaard, W.; Ten Cate, J. M. *Caries Res* **2010**, *44*, 372.

(5)   Turkenburg, J. P.; McAuley, K. E. *Acta crystallographica. Section D, Biological crystallography* **2013**, *69*, 1193.

(6)   Karplus, P. A.; Diederichs, K. *Science (New York, N.Y.)* **2012**, *336*, 1030.

(7)   Race, P. R.; Bentley, M. L.; Melvin, J. A.; Crow, A.; Hughes, R. K.; Smith, W. D.; Sessions, R. B.; Kehoe, M. A.; McCafferty, D. G.; Banfield, M. J. *J Biol Chem* **2009**, *284*, 6924.

(8)   Stein, N. *Journal of Applied Crystallography* **2008**, *41*, 641.

(9)   McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J Appl Crystallogr* **2007**, *40*, 658.

(10)   Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 213.

(11)   Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallographica Section D* **2010**, *66*, 12.

(12)   Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 486.

(13)   Schrodinger, LLC 2010.

(14)   Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. *Acta Crystallographica Section D* **2012**, *68*, 431.