Electronic Supplementary Material (ESI) for Chemical Communications.

Rescuing DNA repair activity by rewiring H-atom transfer pathway in the radical SAM enzyme, Spore Photoproduct Lyase.

Alhosna Benjdia, Korbinian Heil, Andreas Winkler, Thomas Carell, Ilme Schlichting*

*alhosna.benjdia@jouy.inra.fr

Experimental Section

Cloning, expression and purification

The gene coding for the SP lyase from *Geobacillus thermodenitrificans* (GTNG_2348) and the C140A mutant were cloned in our previous study.[1] The C140A/S76C mutant for this study was generated by site-directed mutagenesis (by QuickChange PCR mutagenesis) using the pET-M11-C140A mutant vector as template and the following primers with the modified bases indicated in italics: forward 5'-AAATTTGATAGCTGCAAACCGAGCGCAGA-3', reverse 5'-TCTGCGCTCGGTTTGCAGCTATCAAATTT-3'. The sequence of the newly generated mutant was verified by DNA sequencing. All genes possess a coding sequence for an N-terminal His6-tag allowing their purification by affinity chromatography. The pET-M11-C140A/S76C mutant SP lyase construct was transformed into *Escherichia coli* BL21 (DE3). Expression and purification of the wild-type enzyme, C140A and C140A/S76C mutants were performed as described previously.[1] Briefly, the expression of each protein was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG). The harvested cells were then disrupted by sonication and the supernatants collected after ultracentrifugation. The proteins were purified by Ni-NTA affinity chromatography (Qiagen) followed by a HiTrap Heparin column step. The purity and quantity were checked by SDS-gels and absorbance measurements at 280 nm, respectively.

The iron-sulfur cluster of the purified proteins was then reconstituted as described previously.[1] The proteins were finally concentrated with a 30 000-MWCO spin concentrator (Amicon, Millipore) and stored in aliquots at -80 °C.

Crystallization and structure determination

The reconstituted C140A/S76C mutant SP lyase incubated with Se-SAM was crystallized at 293K in a hanging-drop vapor diffusion setup in the presence of 70 mM octanoyl-N-hydroxyethylglucamide (Hampton Research) and a reservoir solution (200 mM lithium sulfate, 100 mM Tris–HCl pH 9 and 19-27% (wt/vol) PEG 8000) under anaerobic conditions. The crystals were harvested as described previously.[1] The model of SP-bound double mutant SP lyase was obtained from diffraction data of crystals of the C140A/S76C mutant SP lyase soaked in a solution of SP as described for the wild-type enzyme.[1] The C140A/S76C mutant structure was determined by molecular replacement using the wild-type structure as model. XDS and REFMAC were used for data processing and refinement, respectively.

Repair activity of wild-type enzyme and mutants

The activity of the reconstituted wild-type enzyme and the mutants was measured using a 13-mer oligonucleotide containing the dinucleoside SP (5'-CAGCGGT^TGCAGG-3') as substrate.[1] The reconstituted protein (40 µM) was incubated for 4 hours with SAM (3 mM), SP-containing DNA (40 µM), dithiothreitol (5 mM) and sodium dithionite (3 mM) in Tris buffer pH 8, under anaerobic conditions at 25 °C. For the kinetic experiment, we used a 5-fold excess of substrates. Aliquots were taken at different times and the reaction was stopped by flash-freezing

the samples in liquid nitrogen. Samples were diluted in 0.1 M triethylammonium acetate pH 7 (Sigma-Aldrich) for analysis by ultra-fast liquid chromatography (UFLC) (Shimadzu). A reversed-phase C-18 column (Licrospher, 5 µm, 4x250mm from VWR International, Darmstadt, Germany) equilibrated with 98.8% buffer A (0.1 M Triethylammonim acetate pH 7) and 1.2% buffer B (80% acetonitrile, 0.1 M Triethylammonim acetate pH 7) was used. A linear gradient from 1.2 to 25% of buffer B was used at a flow rate of 1 ml min-1. The absorbance was monitored at 260 nm. Standards of SAM, and 5'-deoxyadenosine (Sigma-Aldrich), 6 mer (5'-TGCAGG-3'), 7 mer (5'-CAGCGGT-3') and SP-containing 13 mer were used for the identification of the eluted compounds and we additionally confirmed the nature of products formed during catalysis by mass spectrometry (Figure S3-4).

Mass spectrometry analysis

HPLC purified products eluting at 25.3 and 27.7 min were fractionated and vacuum-dried. Individual samples were resuspended in 0.1 M triethylammonium acetate (Sigma-Aldrich) pH 7 (buffer A) prior to MS analysis using an ultra-fast liquid chromatography setup (UFLC, Shimadzu) coupled to an ESI-MS/MS (maXis ultra-high resolution time-of-flight (TOF) MS, Bruker). A 2 cm C18 reversed-phase guard-column (Discovery Bio C18 packing, Supelco) equilibrated with buffer A was used for additional sample clean-up during elution with a linear gradient from 1.2% to 25% of buffer B at a flow rate of 50 µL min-1. The eluting species were directly infused into the mass spectrometer and analyzed in negative ion mode (capillary voltage - 4.0 kV and capillary temperature - 180 °C). Initially, MS spectra of intact species were acquired in standard operation mode and subsequently species of interest were fragmented by collision induced dissociation (CID). Data were interpreted using the software package Data Analysis (Bruker).

Figure S1: Conserved and similar residues among SP lyases from *Bacilli* and *Clostridia* protein sequences. We performed a Blast search using the SP lyase protein sequence from *G. thermodenitrificans* as query sequence and *Bacilli* and *Clostridia* as chosen organisms. We aligned the 247 sequences with ClustalW2.[2] For clarity, we displayed here only the protein sequence of SP lyases from *G. thermodenitrificans* (*Gt*) and *C. acetobutylicum* ATCC 824 (*Ca*). Fully conserved residues are indicated by asterisk, residues with similar properties are shown by colon and the residues with weak similar properties are displayed by dot. The functionally critical cysteine residue found in *Bacilli* (C140 in *Gt* SP lyase) is highlighted in black. A cysteine residue conserved among in *Clostridium* (C74 in *Ca* SP lyase) is highlighted in grey.

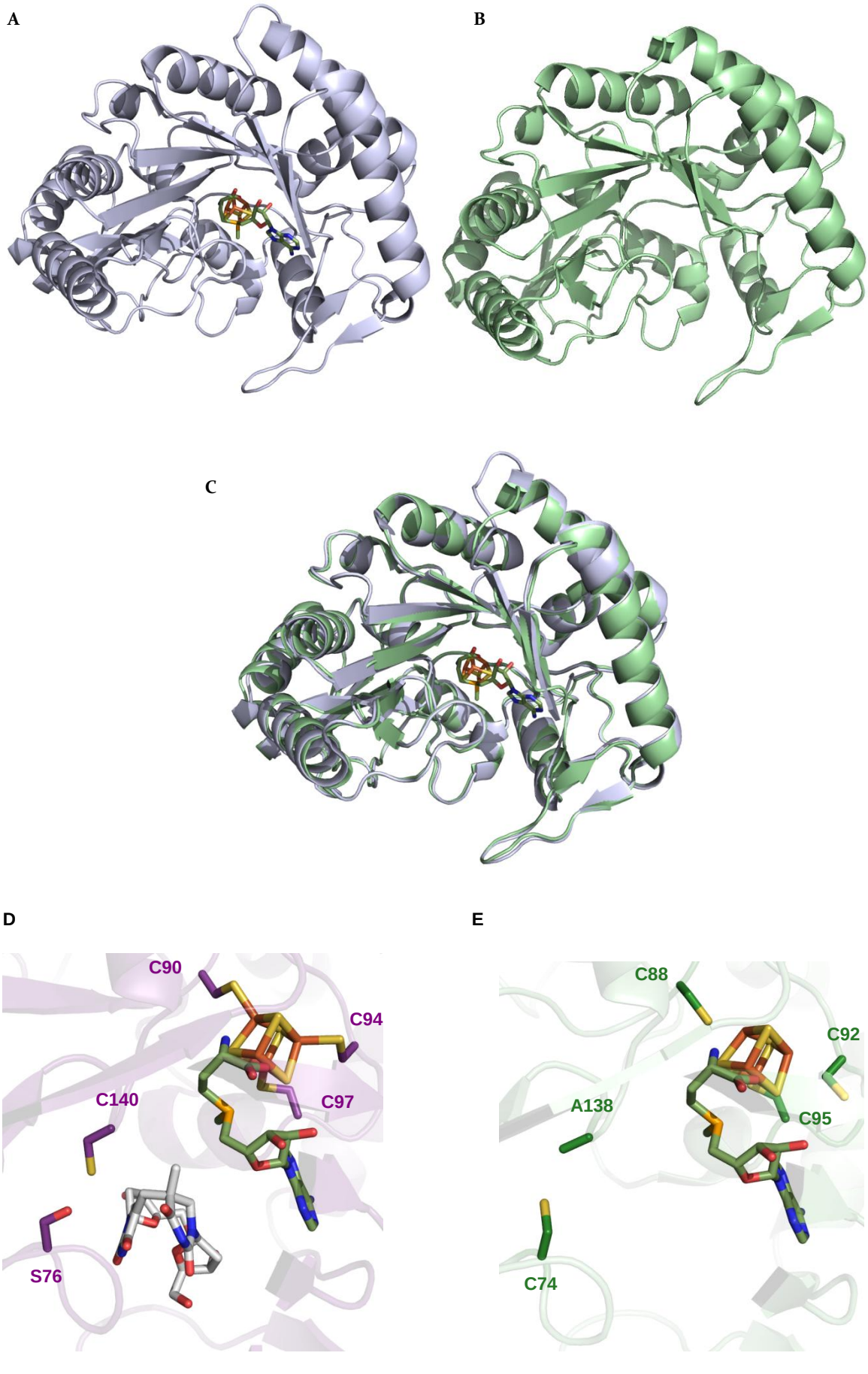

A
B
C
D
C90
C94
C140
C97
S76
E
C88
C92
A138
C95
C74

Figure S2: Secondary structures of SP lyases from *Bacilli* and *Clostridia*. (A) Crystal structure of substrate-free wild-type SP lyase from *G. thermodenitrificans* (PDB code 4FHC). SAM is depicted in green and by atom type. Iron-sulfur cluster is shown in orange and yellow. (B) Structural model of SP lyase from *Clostridium acetobutylicum* ATCC 824 generated by I-TASSER server.[3] (C) Superposition of the crystal structure of substrate-free wild-type SP lyase from *G. thermodenitrificans* (in light purple) and the structural model of SP lyase from *Clostridium acetobutylicum* (in green). The superposition was generated with Coot program using the SSM subroutine on all residues (r.m.s.d. is 0.59 Å).[4] (D) Active site view of substrate-bound wild-type SP lyase from *G. thermodenitrificans* (PDB code 4FHD). Residues are rendered in stick format and colored in purple and by atom type. (E) Active site view of SP lyase model from *Clostridium acetobutylicum*. Residues are colored in green and by atom type.

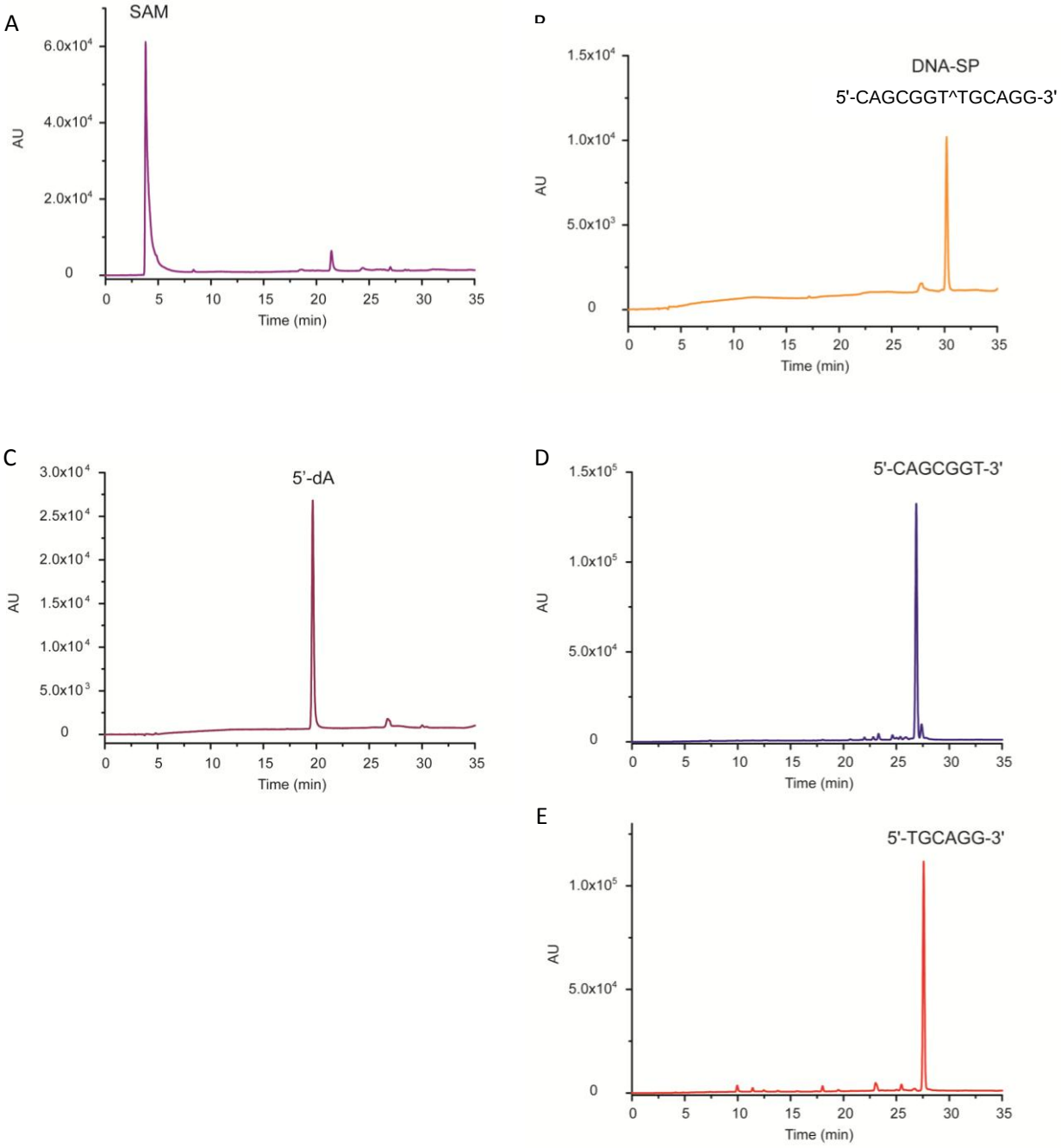
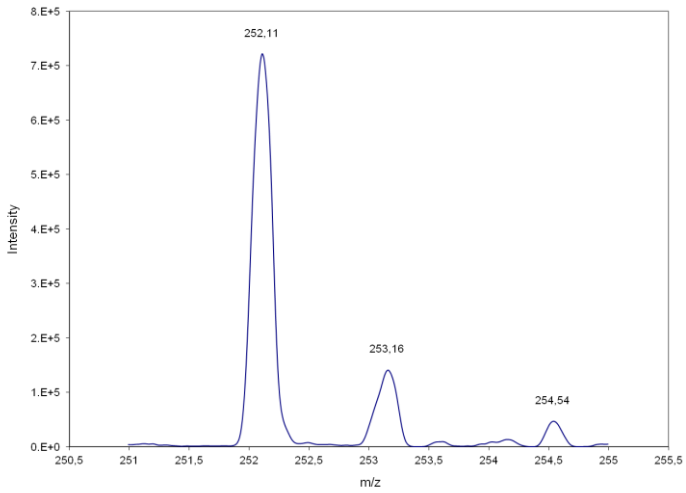
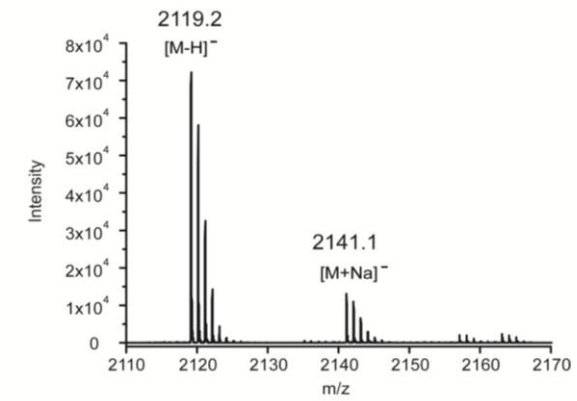


Figure S3: HPLC analysis of standards. (A) *S*-adenosyl-L-methionine (SAM), (B) the SP containing single-stranded DNA substrate (5'- CAGCGGT^TGCAGG -3'), (C) the 5'-deoxyadenosine (5'-dA). As there is no phosphate bridge in the chemically synthesized SP, the repair of the SP containing DNA leads to two oligonucleotides: the 7 mer (5'- CAGCGGT -3') (D) and the 6 mer (5'- TGCAGG -3') (E).

A



B



C

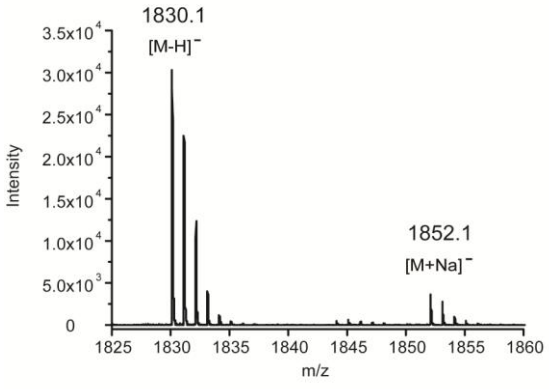


Figure S4: Mass spectrometric analysis of the products formed during catalysis in presence of SP lyase. The products eluted at 19.6 min (A), 27 min (B) and 27.7 min (C) correspond to 5'-dA, the 7 mer (5'-CAGCGGT-3') and the 6 mer (5'- TGCAGG -3'), respectively.

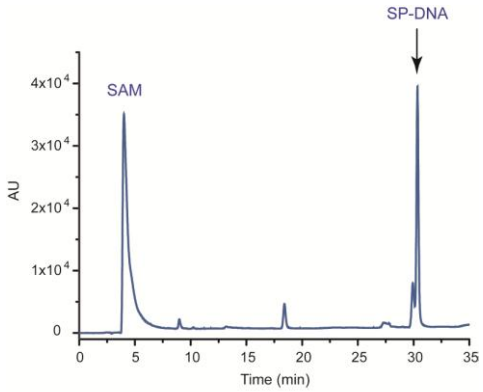


Figure S5: HPLC analysis of the reaction performed with the wild-type SP lyase in absence of an electron donor source. The reconstituted protein (40 µM) was incubated for 1 hour with SAM (3 mM), SP-containing DNA (40 µM), dithiothreitol (5 mM) in Tris buffer pH 8, under anaerobic conditions.

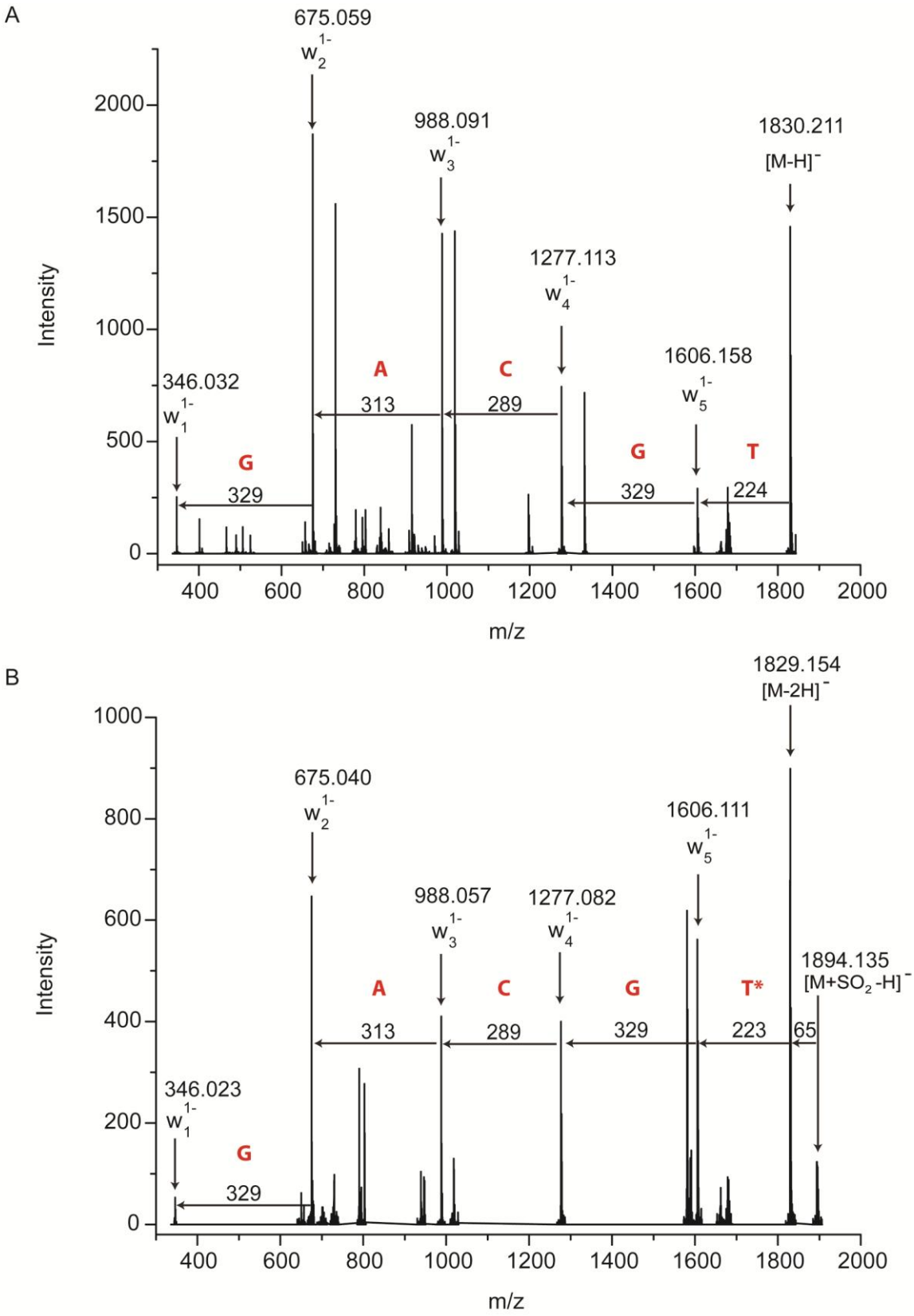
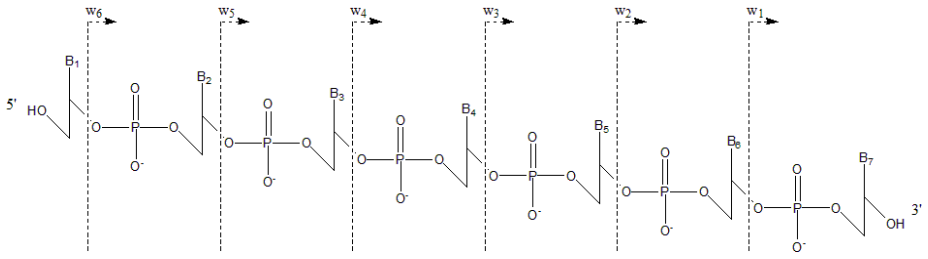


Figure S6: MS/MS analysis of the repaired SP containing oligonucleotide (5'- TGCAGG -3') eluted at 27.7 min (A) and the product eluted at 25.3 min generated only in presence of the C140A mutant (B). $[M-2H]^-$ appears due to preferential cleavage of the sulfinic acid adduct.



ions	Theoretical m/z values	Experimental m/z for the repaired SP containing oligonucleotide (5'- TGCAGG -3') eluted at 27.7 min	Experimental m/z for the product eluted at 25.3 min generated only in presence of the C140A mutant
w_1	346.055	346.032	346.023
w_2	675.107	675.059	675.040
w_3	988.165	988.091	988.057
w_4	1 277.211	1277.113	1277.082
w_5	1 606.264	1606.158	1606.111
$[M-2H]^-$	1829.336	-	1829.154
$[M-H]^-$	1830.344	1830.211	
$[M+SO_2-H]^-$	1894.306	-	1894.135

Table S1: Theoretical and experimental $[M-H]^-$ values obtained after fragmentation of the repaired SP containing oligonucleotide (5'- TGCAGG -3') and the product eluted at 25.3 min generated only in presence of the C140A mutant. The Theoretical m/z values have been simulated for the repaired oligonucleotide (5'- TGCAGG -3') using the massXpert software.[5] (The 1829.154 value originates from the M-2H due to sulfinic acid dissociation).

Table S2: Data collection and refinement statistics

	double mutant SP lyase	double mutant SP lyase + SP
Data collection		
Wavelength (Å)	1.00001	1.07169
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions a, b, c (Å)	54.0, 63, 143.0	53.4, 58.5, 131.9
α, β, γ (Å)	90, 90, 90	90, 90, 90
Resolution (Å)	47.7-2.0	49.5-2.6
High resolution bin	2.0-2.1	2.6-2.7
R merge (%)	8.3 (49.7)	6.8 (45.8)
I / σI	12.7 (3.3)	15.1 (3.3)
Completeness (%)	99.4 (99)	99 (98.9)
Redundancy	5	4.5
Wilson B-factor (Å^2)	28.6	62.2
Refinement		
Resolution (Å)	43.1-2.1	49.5-2.6
No. unique reflections	29294	12449
R_{work} / R_{free}	0.1943 / 0.2295	0.2638/0.3249
No. atoms		
Protein	2809	2801
Fe-S	8	8
SAM	27	27
Water	147	43
SP	-	34
Other ligands*	34	30
Average B-factors (Å^2)		
protein	28.9	45.8
Fe-S	22	39.5
SAM	21.5	34.8
Water	34.2	53.3
SP	-	41.5
Other ligands*	41.5	46.1
R.m.s. deviations		
Bond length (Å)	0.0044	0.0045
Bond angles (°)	0.9408	0.9992

* sulfate, ethylene glycol

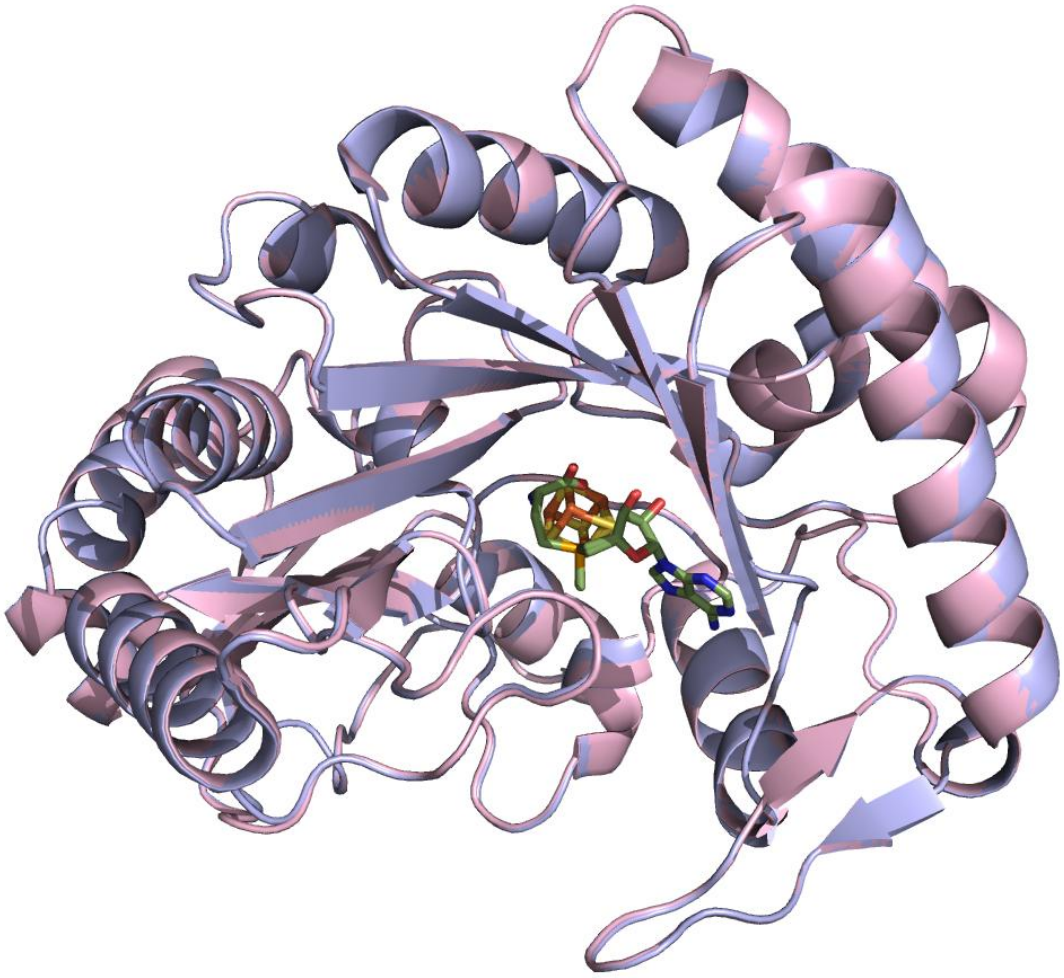


Figure S7: Secondary structure superposition of substrate-free wild-type SP lyase (light purple) (PDB code 4FHC) and the double mutant SP lyase (light pink). The superposition was generated with Coot program using the SSM subroutine on all residues (r.m.s.d. is 0.16 Å).[4]

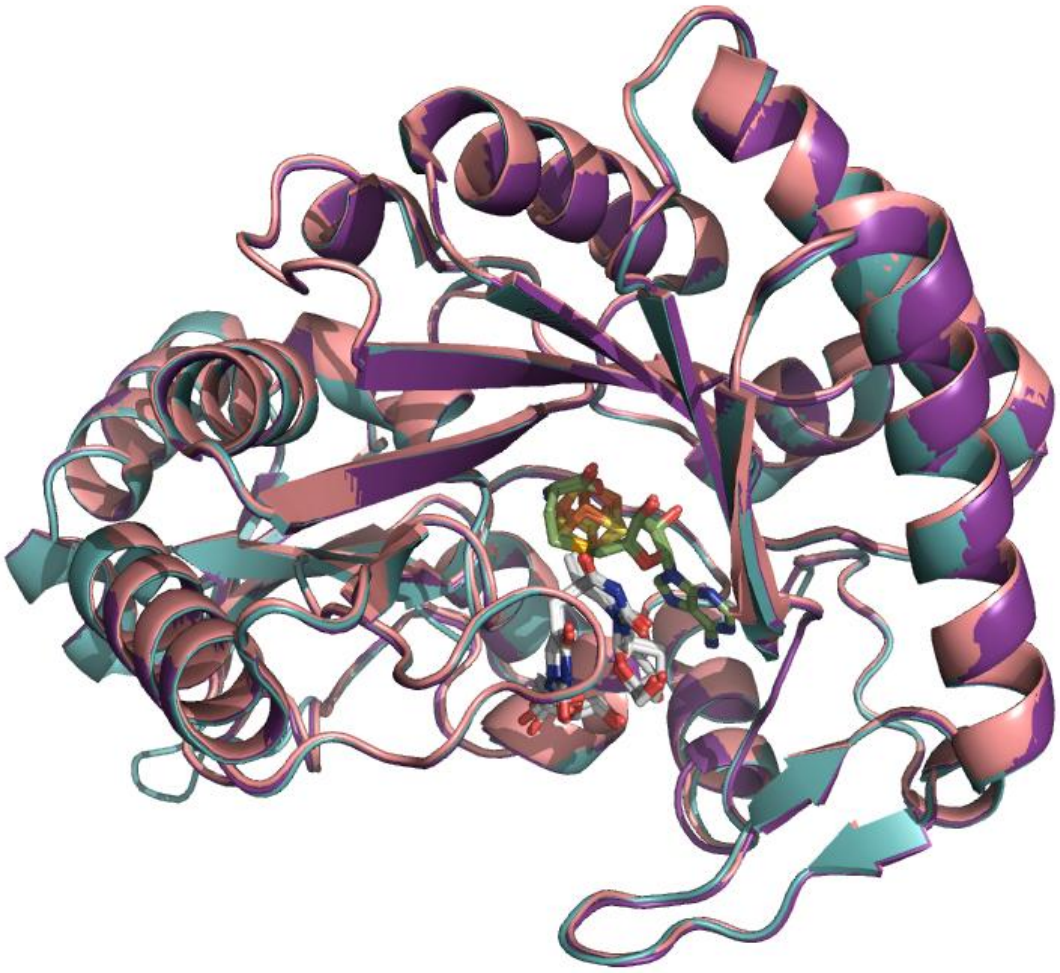
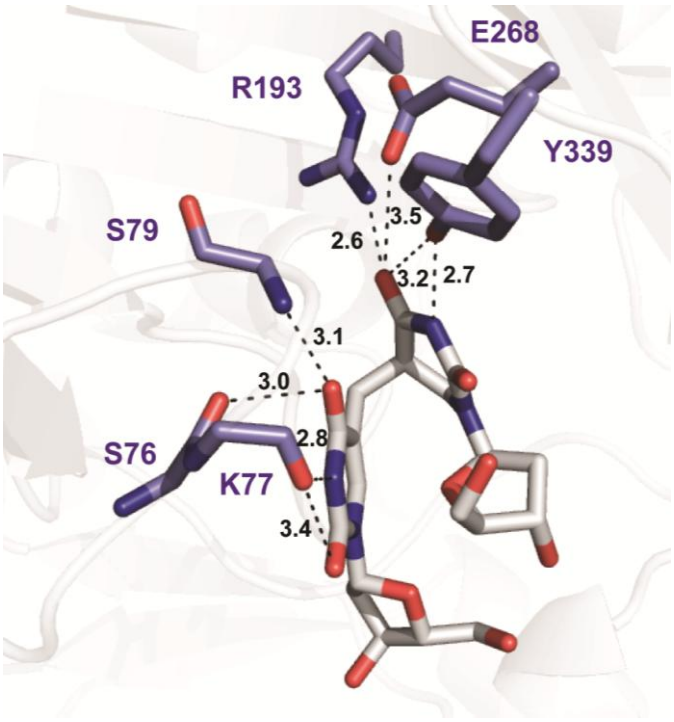


Figure S8: Secondary structure superposition of the substrate-bound wild-type SP lyase (PDB code 4FHD), C140A mutant (PDB code 4FHF) and the double mutant SP lyase. Superposition of the wild-type SP lyase (purple), the C140A (cyan) mutant and the double mutant SP lyase (salmon) were generated with Coot program using the SSM subroutine on all residues (r.m.s.d. is 0.29 Å and 0.21 Å, respectively).[4] The iron-sulfur cluster is depicted in orange (Fe) and yellow (S), SAM is colored in green by atom type and SP is shown in white.

A



B

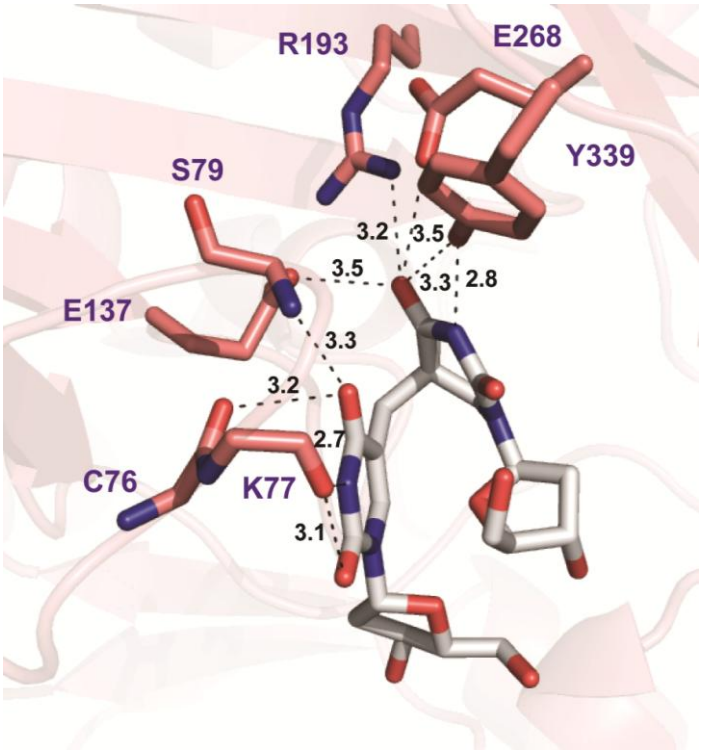


Figure S9: Active site of substrate-bound SP lyases. Side-chain residues of (A) the wild-type SP lyase (PDB code 4FHD) and (B) the double mutant SP lyase in interaction with SP (colored in white) are depicted in purple and salmon, respectively. The distances in Å are indicated in black. This shows that the amino acids involved in the binding of SP superimpose as well (Figure S9).

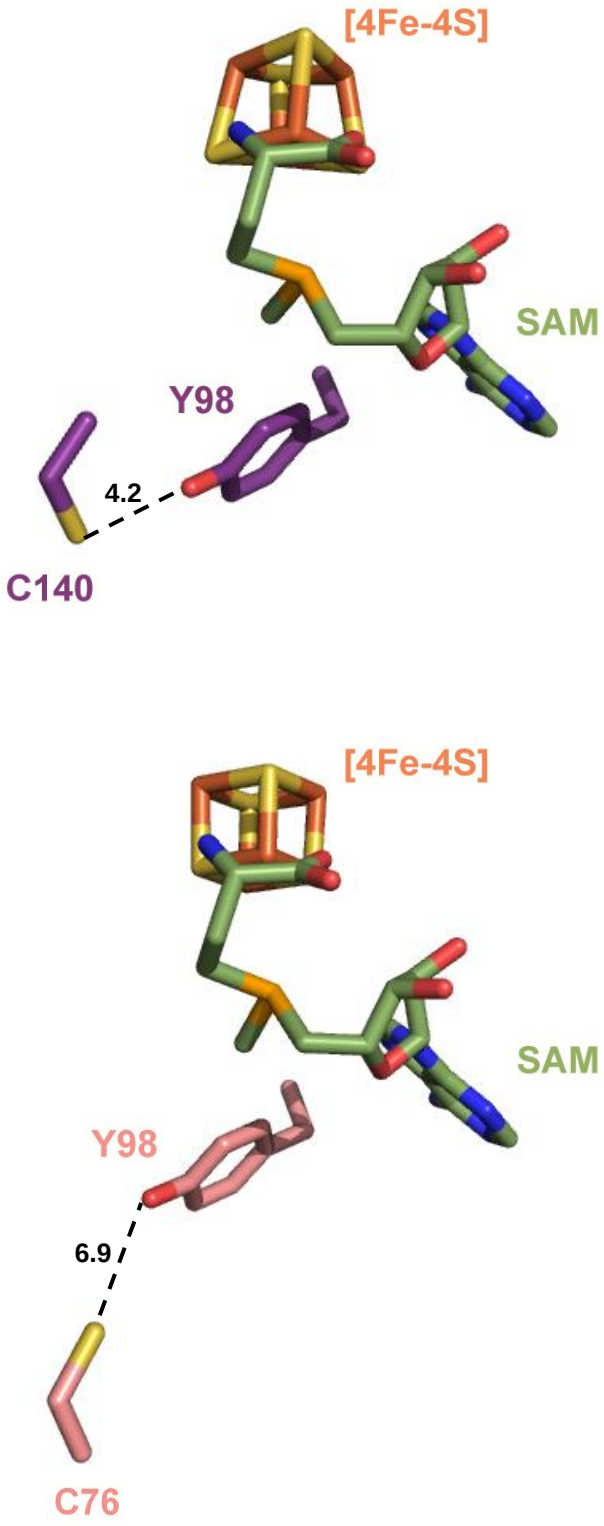


Figure S10: Active sites of substrate-free (A) the WT SP lyase (PDB code 4FHD) and (B) the double mutant SP lyase. Residues are depicted in stick format and colored in purple (WT) and in salmon (the double mutant SP lyase). Distances between the cysteine and the tyrosine residues are rendered in dashed black line.

References

[1] A. Benjdia, K. Heil, T. R. Barends, T. Carell, I. Schlichting, *Nucleic Acids Res* **2012**, *40*, 9308-9318.

[2] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins, *Bioinformatics* **2007**, *23*, 2947-2948.

[3] aA. Roy, A. Kucukural, Y. Zhang, *Nature protocols* **2010**, *5*, 725-738; bY. Zhang, *BMC bioinformatics* **2008**, *9*, 40.

[4] E. Krissinel, K. Henrick, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2256-2268.

[5] F. Rusconi, *Bioinformatics* **2009**, *25*, 2741-2742.