Bacterial strains	Description	Reference
Escherichia coli		
DH5α	F-φ80d <i>lacZ</i> Δ(<i>lacZYA-argF</i>) U169 <i>deoR</i> supE44Δ <i>lacU169</i> (f80 <i>lacZDM15</i>) <i>hsdR17 recA</i> 1 <i>endA1</i> (rk- mk+) supE44gyrA96 thi-	[1]
BL21(DE3)p <i>ly</i> sS	F- $ompT$ hsdS gal (rb- mb+) DE3(Sam7 $\Delta nin5$ lacUV5-T7 Gen1)	[1]
Staphylococcus aureus		
RN4220	restriction negative strain/MSSA cloning intermediate derived from 8325-4	[2]
COL	Archaic HA-MRSA strain	[3]
COL-∆ <i>aldA</i>	COL aldA mutant	This study
COL-∆aldA::pRB473-aldA	COL aldA complemented strain	This study
COL-∆aldA∷pRB473-aldAC279S	COL <i>aldAC279S</i> complemented strain	This study
COL-∆sigB	COL sigB mutant	[4]
Staphylococcus phage 80		[5]
Discusida		Deference

Table S1: Bacterial strains and plasmids

Plasmids	asmids Description	
pET11b	E. coli expression plasmid	Novagen
pRB473	pRB373-derivative, <i>E. coli/ S. aureus</i> shuttle vector, Amp ^r , Cm ^r	[6]
pRB473-XylR	pRB373-derivative, <i>E. coli/ S. aureus</i> shuttle vector, containing xylose-inducible P_{Xyl} promoter Amp ^r , Cm ^r	[7]
рМАD	Shuttle vector for allelic exchange in <i>S. aureus</i> , Amp ^r , Em ^r	[8]
pET11b- <i>aldA</i>	pET11b overexpressing His ₆ -AldA	This study
pET11b- <i>aldAC27</i> 9S	pET11b overexpressing His ₆ -AldAC279S	This study
pRB473- <i>aldA</i>	pRB473 expressing <i>aldA</i> under P _{xyl}	This study
pRB473- <i>aldAC279S</i>	pRB473 expressing aldAC279S under P _{Xyl}	This study
pMAD-∆ <i>aldA</i>	pMAD-∆ <i>aldA</i> for <i>aldA</i> mutant construction	This study

Table S2.	Oligonucleotide	primers
-----------	-----------------	---------

Primer name	Sequence (5' to 3')					
aldA-for-Ndel	GGAATTC <u>CATATG</u> AGAGACTACACAAAGCAATAC					
aldA-rev-BamHI	CGC <u>GGATCC</u> TTAGTGATGGTGATGGTGATGTTTAAAATATCCAGCT					
	ATAGATTTCAC					
aldA-pMAD-up-for	CGC <u>AGATCT</u> ATCTCTCAAGATGTAATGACTG					
aldA-pMAD-up-rev	ATCAAATTTATTTAAAATATCCAGCCTCATTATTCTTTCACTCCTCA					
aldA-pMAD-do-for	TGAGGAGTGAAAGAATAATGAGGCTGGATATTTTAAATAAA					
aldA-pMAD-do-rev	CCA <u>GTCGAC</u> AATAGTTCACTGTTCTTAAAAAAC					
aldA-pRB-for-BamHI	TAG <u>GGATCC</u> TCATTCTTGAGGAGTGAAAGAATAATGAGAGACTACA					
	CAAAGC					
aldA-pRB-rev-KpnI	CTC <u>GGTACC</u> TTATTTAAAATATCCAGCTATAGATTTC					
aldA-C279S-rev	TGTACCAGCTGT AGA TACTTGACCAGTATTATT					
aldA-C279S-for	AATAATACTGGTCAAGTA TCT ACAGCTGGTACA					
aldA-for	CCGTGGAACTTCCCTACAAA					
aldA-rev	CTAATACGACTCACTATAGGGAGATGGCTTTCTACCTGCTTCGT					

Restriction sites are underlined and bold bases indicate point mutations

Table S3. Structural comparison of saAldA (PDB 3TY7) with other aldehyde dehydrogenases.

PDB ID	Enzyme	R.m.s.d. [Å]	Cα atoms aligned	Sequence identity / similarity [%]
3144	Bartonella henselae aldehyde dehydrogenase (bhADH)	1.39	438	41 / 61
4126	Pseudomonas fluorescens 2-aminomuconate-6-semialdehyde dehydrogenase (pfAMSDH)	1.39	428	35 / 53
2WME	Pseudomonas aeruginosa betaine aldehyde dehydrogenase (paBADH)	1.54	434	38 / 57
1WND	<i>Escherichia coli</i> aldehyde dehydrogenase (<i>ec</i> ADH)	1.50	423	34 / 51
418P	<i>Zea mays</i> aminoaldehyde dehydrogenase (<i>zm</i> AMADH)	1.40	435	36 / 52
4A0M	<i>Spinacia oleracea</i> betaine aldehyde dehydrogenase (<i>so</i> BADH)	1.40	433	38 / 57
3IWK	<i>Pisum sativum</i> aminoaldehyde dehydrogenase (<i>ps</i> AMADH)	1.40	432	38 / 57
4NU9	Staphylococcus aureus betaine aldehyde dehydrogenase (saBADH)	1.40	427	39 / 58
4I9B	Solanum lycopersicum aminoaldehyde dehydrogenase (s/AMADH)	1.43	434	39 / 56
1005	Homo sapiens mitochondrial aldehyde dehydrogenase (hsALDH2)	1.43	426	35 / 52



Figure S1: Growth curves of *S. aureus* **COL wild type and the** *aldA* **deletion mutant after exposure to formaldehyde (A,B) and methylglyoxal (C,D).** The *S. aureus* COL wild type and the *aldA* deletion mutant were grown in RPMI medium and exposed to sub-lethal concentrations of 0.5 mM and 0.75 mM formaldehyde **(A,B)**, 0.5 mM and 0.8 mM methylglyoxal **(C,D)** at an OD₅₄₀ of 0.5 during the log phase. No growth phenotype of the *aldA* mutant was detected under aldehyde stress.



Figure S2. AldA is not required for the survival of *S. aureus* under formaldehyde stress. For the survival phenotype assays, *S. aureus* COL wild-type (WT), the $\Delta aldA$ deletion mutant (A) and the *aldA* and *aldAC279S* complemented $\Delta aldA$ mutants ($\Delta aldA$ pRB473*aldA* and $\Delta aldA$ pRB473*aldAC279S*) (B) were grown in RPMI until an OD₅₀₀ of 0.5 and treated with 2 mM formaldehyde stress. Survival assays were performed by spotting 10 µl of serial dilutions after 1 and 3 hours of NaOCI exposure onto LB agar plates.



Figure S3: Transcriptional induction of *aldA* under formaldehyde, methylglyoxal, NaOCI, diamide, H_2O_2 and MHQ stress in the *S. aureus* COL wild type, $\Delta aldA$ mutant, *aldA* and *aldAC279S* complemented $\Delta aldA$ mutant strains. RNA was isolated from *S. aureus* COL wild type under control, 0.75 mM formaldehyde (FA), 0.5 mM methylglyoxal (MG), 1 mM NaOCI, 2 mM diamide, 10 mM H_2O_2 and 50 μ M methylhydroquinone (MHQ) stress conditions. RNA samples were also prepared from the $\Delta aldA$ mutant as well as from the *aldA* or *aldAC279S* complemented strains under control conditions in the presence of 1% xylose. The RNA samples were subjected to Northern blot analysis for *aldA* (SACOL2114) transcription. The successful *aldA* and *aldAC279S* complementations are indicated. The methylene blue stain is the RNA loading control showing the abundant 16S and 23S rRNAs.



Figure S4. Cys279 is essential for AldA activity towards methylglyoxal and formaldehyde oxidation. Purified AldAC279S shows no significant activity for oxidation of methylglyoxal (MG) **(A)** and formaldehyde (FA) **(B)** and *in vitro*. Reduced AldAC279S protein (2.5 μM) was incubated with different concentrations of methylglyoxal **(A)** and formaldehyde **(B)** in reaction buffer (100 mM Tris HCI, 1.25 mM EDTA, pH 7.5). The oxidation of the aldehydes was measured in the presence of NAD⁺ as coenzyme and NADH generation was monitored at 340 nm using a spectrophotometer. The results are from 3 replicate experiments. Error bars represent the SEM.

					-			
saAidA bhADH prAMSDH zmAMADH soBADH saBADH saBADH siAMADH hALDH2 ecALDH paBADH	1 MNTLPSOVM 1 MASPAM 1 MASPAM 1 MAIT- 1 MELLKH 1 MANRNVP- 1 MLRAARFGPRLG 1 MARFEE	RTNIG RRLLSAAATQAVPAPNQQ	11 	0 20 CEWVESNSNET GLWDDPSTPHD CNFVTSASSFA GEWREPIKKNR GEWREPAGGRR GEWVESANKNT GEWVESANKNT GEWVESANKNT GEWVESANKNT GEVVESASSGAT	30 I L V I NP ATEE I V I OPSTEL N - INPVNGK I PVI NPSTEL IPVI NPSTEL IPVI NPSTEL IPI NPATEE GPVVNATGO FET NPANGE	40 A G KVAKGNKA CAVISLGSTR I BOVFEADAK HIGEI PAGTAE HIGDI PAATAE HIGDI PAATAE I GDI PAATAE I GDI PAATAE VI COVAEGOKE VL EI AEASAE VL AKVQRASRE	50 DVDKAVE AA DDVYI DADKAINAKKAF OVNEAVVAARAAE DVDAAVAARAEI DVDLAVDAAKRAI DVEVAVVAARAEI DVDLAVDAAKAI DVELAVRAARAEI DVDIAVRAAAAAF DVDKAVKAARAAF	60 LEFRHTSV 63 QTWKTTSP 63 KGPWGKLSV 77 KRNRGROWARAPG 73 SRNNWSATSG 68 SRKNGROWSAASG 71 ESGEWSQETA 69 ARDDWGSTTG 71 QLGSPWRMDA 89 AEWGQTTP 60 KVWAAMTA 65
SƏAIdA bhADH prAMSDH zmAMADH soBADH psAMADH saBADH siAMADH hALDH2 ecALDH paBADH	70 64 KERQAL DKIVKE 64 HERLGFVEKILE 178 QDRAAL IHKIADG 74 AVRAKYLRAIAAK 69 AHRATYLRAIAAK 72 SLRARYLRAIAAK 70 ET GKKVRAIADK 70 ET GKKVRAIADK 99 SHRGRLNRIADI 61 KVRAECLLKLADV 66 MQRSRILRRAVDI	80 90 YENRKDDIVQAITDELGA YEKRSSDMAKTISMEMGA VIERKPEFVAAUDTGR VIERKPELAKLEALDCGK IKEKKDELGKLESIDSGK IKEKKSVLATLESLDSGK VLEKKSVLATLESLDSGK FERDRTYLAALETLDNGK LRENRGVFAELESRNCGK LRERNDELAALETLDTGK	100 PLSLSERVHYG FIDMALNAQTA VHQARTDIP PYDEA-AWDMD FDEA-VLDID PLEEA-LADLD TLES-YADMD TLES-YADMD PLHSAFNDEIP PLAETRSVDIV	110 MGLNHEVAARD TGSSHIRNFIK RAIANFTFAD DVASCFEYFAG DVACFEYFAG DVACFEYFAG DVACFEYFAG NVLKCLRYAG AIVDVERFFAG TGADVLEYYAG	120 ALD NYE E AYK EFSJ LAKTSHTDLFI QAEALDKRONS LAELDSKOK, LAELDSKOK, LAELDSROM ACS	130 FEER RG DD - GOEAL IEGNEQ MSTS DGSGAL SPVSLPMETFK APVILPMETFK APVILPMETFK MIDSPIPDTE FPVNLNSDSYK MIDSPIPDTE FPVNLNSDSYK - AAGEYLEGHT OIPLR - ETSF	140 - LVVKEAIGVSGL AILHVDAIGVVGL NYTVRKEPIGVVGL SVILKEPIGVVGL SVILKEPIGVVAL SVILKEPIGVVAC SVIVLREPLGVVGL SVIREPLGVVGS SVIREPLGVVAS VYTRREPLGVVAS	150 IT PWN FP TNQTSL 15 IT PWN WMNQYTL 15 IS PWN PLLAFTW 17 IS PWN YPLLMÄTW 17 IS PWN YPLLMÄTW 16 IT PWN YPLLMÄTW 16 IT PWN YPLLMÄTW 17 IT PWN YPLLMÄTW 17 I I PWN FPLLMÄTW 17 I I PWN FPLLMÄTW 15 I A DWN YPLMÄÄW 15 I GAWN YPVQ I ALW 16
saAldA bhADH prAMSDH zmAMADH soBADH psAMADH saAMADH slAMADH hALDH2 ecALDH paBADH	160 156 KLAAARAAGSPV 100 KVIPALAGGTMV 178 KVAPALAGGTMV 178 KAPALAAGGTAV 178 KAPALAAGGTAV 171 KIPALAAGGAT 171 KVAPALAAGGAL 171 KVAPALAAGGAL 171 KVAPALAAGANVV 185 KLAPALAAGNVV 185 KLAPALAAGNVV 185 KLAPALAAGNVV	170 180 LKPSE FPFAAVIAE LKPSE IAPLSAMLFAE AKPSE ESPSATLAE LKPSE LASVTCLEGE LKPSE LYNNLI LKPSE LYNNLI LKPSE LYNNLI LKPSE LYNLI LKPSE LYNLI	190 DKVGVPKGVF DEAALPSGVF HDAGVPPGVF KEVGLPSGVF NEVGLPSGVF KEVGLPKGVL KEVGLPKGVL KEVGLPKGV KELGLPSGAL KELGLPSGAL KDI-FPAGVI KDI-FPAGVI	200 L V NG DG AG - VG L I NG DG AN - VG L I NG FG KD SAG I V TG LG PD - AG I V TG LG PD - AG I V TG LG HE - AG I L LG AG SE - VG I L TG LG PE - AG I V PG FG PT - AG I L FG RG KT - VG V L TG SG RE - VG	210 NPLSAHPKVR SYLSAHPDLGI EFLTQHPGIS/ APLSAHPDVDI ASLASHPDVDI OVMSGHKEVDI GPLASHPHVDI DPLTGHPKVR QWLTEHPLIEI	220 MISTICSOPIC ALTIGSTA CVA TOSFEIG CLATOSFEIG CLATOSATO CLATOSAT	230 240 SKIMEKAKK DFK KDISKNSN-TLK KKIMASAP-MVK SKIMKAVAD-GVK SKIMTAAQ-LVK KHIMKNAAN-NVT SKIMTAAQ-LVK KHISHTAS-SIK KKVMASASSSLK) 250 KVSL LGGKSPV 25 KVSL LGGKSAN 25 VSFLLGGKSAN 25 VTL LGGKSPV 26 VTL LGGKSPV 26 VTL LGGKSPV 28 VTL LGGKSPV 28 RTHMLLGGKSPV 28 RTHMLLGGKSPV 25 VTL LGGKSPV 25 RTHMLLGGKSPV 25 VTMLLGGKSPV 25 VTMLCGKSPV 25
səAldA bhADH prAMSDH zmAMADH soBADH psAMADH srBADH srAMADH hALDH2 ecALDH paBADH	260 254 ULDOVD - IKEAAK 259 IFADAD - LDAAR 277 VFADAD - LDAAR 277 VFADAD - LDAAR 276 VFDVD - IDKVE 289 VFDVD - IDKVE 289 VFDVD - LDKVE 289 VFDVD - LAVVE 289 VFDVD - LAVVE 289 VFDVD - IEAVVE 255 VFDDAD - IEAVVE 261 IFPDAD - LDRAAD	270 280 ATTGKVVNNTGVTTAGT GVR-HCFVNSGOS NAPT GVLRSFTNSGOVLCSE WTLFGCFWTNGOISATS WTLFGCFWTNGOISATS MTVFGCFFTNGOISATS MTLFGIFANTGQVSATS GVRTFGYVNAGODTAAC GVRTFGYVNAGODTAAC	200 RVLVPNKIKA RVVVHRSIFDE RLLVHRSIFDE RLLVHESIAAE RLLVHESIAAE RLIVHESIAAE RLIVHESIAAE RLIVHESIAVE RILVORNIASA RILVORNIASA RIVORNIASA	300 LAELKEGF5G FVSGLKVEAEF FNEMVAWAKN FVDKLVKWTKN FVDKLVKWEK FMDRLLKWTKN EVERVARAKS LVEKLGAAVAT	310 VRVGNEREDG ITQVGPGHQTG ITQVGPGHQTG ITVSDPLEEG ITSDPLEEG ITSDPLEEG ITSDPLEED ITSDPLEED VSONPFDSK LKSGAPDDES ITRLGDPQDEN	320 TO VGP II SKKO NHIGP VVS KEO VMG PLI SHGH CRL GP VS EGO CRL GP VS KGO CRL GP VS SKGO TEMGP VS STEH CKL GP VVS AGO TEL GP LSS LAH TNF GPL VS FPH	330 3 FDQVQNYINKGIE YDKIODLOSGID RDKVLSYYRLÄVD YEKIKKFISNAKS YDKIMKFISNAKS YKKVLNCISSAKS YKKVLNCISSAKS YKKVLKFISNAKS FKKILQVINTGKO LERVGKAVEEAKA MESVLGYIESGKA	40 350 EG RELFYGGPGK 55 EG ATLVTGGTGL 35 EG ATLVTGGVF 37 GG AT ILTGGVF 36 EG AT ILTGGVF 36 EG AT ILTGGRF 36 EG AT ILCGGRF 36 EG AT ILCGGGI A 39 TGHIKVITGGEKR 35 GK ARLCGGERV 35
saAldA bhADH prAMSDH zmAMADH soBADH paAMADH saBADH siAMADH hALDH2 eeALDH paBADH	360 352 PEG EEK YFAR 355 PMG - MERGYVR 355 KFNDERDQAYVR 369 A HE KKOYYE 364 E HE KKOYYE 367 E HE KKOYYE 368 Q HE KKOYYE 388 Q - HE KKOYYE 388 Q ROYEYA 392 AD ROYEYA 395 TDG - AFGKGAYVA	370 380 PT IF INVDNONT ACE I VFADVKPHNRIFRE I TI WTGLSDKARCVT EI I ITOI ITSNE WRE V VTUTOI STSNO WKE V VTUTOLSTSNO WKE V VTUTOCDTSNE VCE V VTUTOCDTSNE VCE V VFGDVADGNT AKE I PTVFDCRDDNT VRE I	BOUND STATES		410 TEYGLAGYUIO SNYGLAGAV TOYGLAGAV TOYGLAGAV TEYGLAAVF THYGLGSAVM SIYGLAGAV STYGLAAAVF TKYGLGAAVM STYGLAAAVF TEYGLAASVW	420 GKD KETLHKVA SO PRSKCR FIA SO PRSKCR FIA SSO PRERCORLS SG PRERCORLS SKD LERCER IT SND LERCER IT SND LERCER FT TKD UCANYLS TKD UCANYLS TKD UCANARYS TCD LARAHRAI	430 RS I EA T VE I NEAL AQ VR 36 MV EV NG HI RQ I HV CL V WV T W' EE I DA C I NV N CS' KAL EV CA V WV N CS' KAL QA L V WI N CA' NKL KL T V WI N DFI KAFQT I I WI N CS' QAL QA CT V WV N CYI ARL QY CC T WV T HI H RL EAG I C WI N TW	440 GRK - PDL FF GQ YK 44 EL P - GGS YF GQ VK 45 YL RDL RT FF GQ VK 47 QPCF CQ A WGQ I K 46 QPCF VQ A WGQ I K 46 QPCF VQ A WGQ Y 46 HPY FAQ A WGQ Y 46 DVF GQ AS FF GQ Y 48 FML VSEM HGQ K 44 GES PAEME VGQ YK 45
saAidA bhADH prAMSDH zmAMADH soBADH saBADH saBADH slAMADH hALDH2 ocALDH	450 460 449 Q SG L G R E WS D Y 452 F SG R A R EG L WS 475 L SG L G R E C R F S 466 R SG F O R L C E G 461 R SG F O R L C E WG 464 R SG F G R L C E WG 464 R SG F G R L C E WG 465 Q SG R L G R L C K E G 465 Q SG R L G R L C K E G 466 Q SG R L C E Y G 469 A SG R L C Y G R W S L Y G 469 A SG R C V G R W S L Y G	470 EEFLEVKSIAGYFK KEFLDTKAIS-YW DOYLSVKOVTEYISDEPW ONYLNIKOVTQJISDEPW ENYLSVKOVTRYTSDEPW ENYLSVKOVTRYTSDEPW ENYLSVKOVTRYTSDEP ENYLSVKOVTYYKVPQKNS EDYTYVKPVPQKNS	GWYQSPSKL GWYKSP- GWYQPPSKL GWYSKS AFYKSPSKN AFYKSPSKN	175 176 100 105 197 103 103 104 104 104 117 174 190				

Figure S5. Sequence alignment of *sa*AldA and other ADHs (as in Table S3). Amino acid sequences were aligned using Clustal Ω [9] and presented using Jalview [10]. Intensity of the blue color gradient is based on sequence identities and similarities. The highly conserved residues are highlighted with colors (catalytic cysteine and glutamate residues – red, glutamate and lysine residues involved in catalysis – light pink; cation-binding residues – yellow). The secondary structure elements above the alignment correspond to the apo-*sa*AldA structure and are colored by domains (CoBD – blue; SID – green; CD – magenta). Numbering above the alignment is according to *sa*AldA.



Figure S6. Plot of the dihedral distribution of N-CA-CB-SG dihedral (rotation around the CA-CB bond) of AldA Cys279 (A) and plot of the dihedral angle of AldA Cys279 in the apo- and holo-enzymes as the function of the simulation time (B). The plots show that Cys279 in the apo-enzyme (black) has very different dihedral propensity than in the holo-enzyme complex with NAD⁺ (magenta), explaining the preference of Cys279 to form the BSH complex in the apo-enzyme in the "resting" (Q2) position and in the holo-enzyme in the "attacking" (Q1) positions.



Figure S7. S-bacillithiolation of AldA Cys279 does not require major structural changes as revealed by MD simulations. Root-mean-square deviation (RMSD) of AldA protein backbone (A) and per-residue average root mean square fluctuation (RMSF) (B) during 50 ns of MD simulations of AldA apo-enzyme (black), AldA holo-enzyme complex with BSH in "attacking" (Q1) position (red), and AldA apo-enzyme complexes with BSH in "resting" (Q2) position (green). S-bacillithiolation resulted in little change in the backbone flexibility of AldA between the complexes of BSH with the apo-enzyme (Q2), the holo-enzyme (Q1) and the apo-enzyme without BSH.

REFERENCES

[1] Studier, F. W.; Moffatt, B. A. Use of Bacteriophage-T7 RNA-Polymerase to Direct Selective High-Level Expression of Cloned Genes. *J Mol Biol* **189**:113-130; 1986.

[2] Kreiswirth, B. N.; Lofdahl, S.; Betley, M. J.; O'Reilly, M.; Schlievert, P. M.; Bergdoll, M. S.; Novick, R. P. The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* **305**:709-712; 1983.

[3] Shafer, W. M.; Iandolo, J. J. Genetics of staphylococcal enterotoxin B in methicillin-resistant isolates of *Staphylococcus aureus*. *Infect Immun* **25**:902-911; 1979.

[4] Pane-Farre, J.; Jonas, B.; Förstner, K.; Engelmann, S.; Hecker, M. The sigmaB regulon in *Staphylococcus aureus* and its regulation. *Int J Med Microbiol* **296**:237-258; 2006.

[5] Rosenblum, E. D.; Tyrone, S. Serology, Density, and Morphology of Staphylococcal Phages. *J Bact* **88**:1737-1742; 1964.

[6] Bruckner, R.; Wagner, E.; Gotz, F. Characterization of a sucrase gene from *Staphylococcus xylosus*. *J Bact* **175**: 851-857; 1993.

[7] Pöther, D. C.; Gierok, P.; Harms, M.; Mostertz, J.; Hochgräfe, F.; Antelmann, H.; Hamilton, C. J.; Borovok, I.; Lalk, M.; Aharonowitz, Y.; Hecker, M. Distribution and infection-related functions of bacillithiol in *Staphylococcus aureus*. *Int J Med Microbiol* **303**:114-123; 2013.

[8] Arnaud, M.; Chastanet, A.; Debarbouille, M. New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70**:6887-6891; 2004.

[9] Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Sys Biol* **7**:539; 2011.

[10] Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**:1189-1191; 2009.