Chemistry Europe
European Chemical Societies Publishing

Engineered Sortases in Peptide and Protein Chemistry

Christian Freund*[a] and Dirk Schwarzer*[b]

The transpeptidase sortase A of *Staphylococcus aureus* (Sa-SrtA) is a valuable tool in protein chemistry. The native enzyme anchors surface proteins containing a highly conserved LPxTG sorting motif to a terminal glycine residue of the peptidoglycan layer in Gram-positive bacteria. This reaction is exploited for sortase-mediated ligation (SML), allowing the site-specific linkage of synthetic peptides and recombinant proteins by a native peptide bond. However, the moderate catalytic efficiency and specificity of Sa-SrtA fueled the development of new biocatalysts for SML, including the screening of sortase A variants form microorganisms other than *S. aureus* and the directed protein evolution of the Sa-SrtA enzyme itself. Novel display platforms and screening formats were developed to isolate sortases with altered properties from mutant libraries. This yielded sortases with strongly enhanced catalytic activity and enzymes recognizing new sorting motifs as substrates. This minireview focuses on recent advances in the field of directed sortase evolution and applications of these tailor-made enzymes in biochemistry.

1. Introduction

Bacterial transpeptidases have gained considerable attention as promising tools in modern biochemistry and biotechnology. Sortase A of *Staphylococcus aureus* is the most widely used class of transpeptidase employed for bioconjugation and seminsynthesis of proteins.[1] In bacteria, sortase A is inserted into the bacterial cell wall and cross-bridges surface proteins to the peptidoglycan, facilitating bacterial evasion of the host's immune system.[2] The enzyme recognizes a sorting motif LPxTG, (with x = any amino acid) in the C-terminal region of the surface protein, and cleaves this sequence at the threonine residues.[3] Upon cleavage, the peptide chain downstream of the threonine residue is released as leaving group, while the upstream fragment remains tethered to the active site cysteine as threonine-thioester (Figure 1a).[4] In the following, the enzyme is deacylated by ligating the sortase-bound thioester to the N-terminal amino group of the pentaglycine moiety located at the peptidoglycan cross-bridge. Two features render sortases attractive tools for protein chemistry: The membrane anchor can be removed from the enzyme without loss of transpeptidase activity, and resulting N-terminal truncated enzymes are soluble.[5] Furthermore, the LPxTG motif and N-terminal glycine moieties are sufficient as sortase A substrates, with two glycines minimally required for full activity of the wild-type enzyme. Upon incorporation of the sorting motif and N-terminal glycine into two separate peptides or proteins, sortase A is capable of ligating the two fragments.[6] This methodology, also referred to as sortase-mediated ligation (SML) or "sortagging," is widely applied in modern biochemistry (Figure 1a).[7]

Up to date sortase A of *S. aureus* lacking the N-terminal membrane anchor (Sa-SrtA) is the most widely used enzyme for SML. However, SML catalyzed by Sa-SrtA also suffers from shortcomings hampering broader applications. The reversibility of SML takes a toll on ligation yields, but engineering of sortase substrates by installation of depsipeptide bonds into the sorting motif, inducing secondary structural elements or metal-assisted SML provide efficient strategies for addressing this problem.[8] The main downsides of the catalyst – wild-type Sa-SrtA – are the high K_M value for the sorting motif, poor catalytic activity, a narrow specificity range for the LPxTG sorting motif, and the requirement of calcium as a cofactor.[3] Researchers have recently started addressing these issues by investigating previously uncharacterized sortases from other microorganisms and by directed evolution of Sa-SrtA. While these developments are still ongoing, successful engineering of Sa-SrtA has been reported over the last ten years, representing the central focus of this minireview.

[a] *Prof. Dr. C. Freund*
Freie Universität Berlin, Institute of Chemistry and Biochemistry
Thielallee 63, 14195 Berlin (Germany)
E-mail: chfreund@zedat.fu-berlin.de
[b] *Prof. Dr. D. Schwarzer*
University of Tübingen, Interfaculty Institute of Biochemistry (IFIB)
Auf der Morgenstelle 34, 72076 Tübingen (Germany)
E-mail: dirk.schwarzer@uni-tuebingen.de

This article is part of a joint Special Collection with the Journal of Peptide Science on SPP 1623: Chemoselective reactions for the synthesis and application of functional proteins. Please see our homepage for more articles in the collection.

2. Natural Sortase Variants

Sa-SrtA served as the model enzyme for uncovering the function and activity of sortases in bacteria and belongs to the class A sortases. In addition to the class A enzymes, sortases of classes B-F have been identified, which differ in function and specificity from Sa-SrtA.[9] For example, sortase B of *S. aureus* (Sa-SrtB) recognizes an NPxTG motif and is involved in bacterial iron uptake.[10] However, several sortases of classes B-F show limited activity *in vitro* and are not commonly used for SML. Consequently, research on new catalysts of SML has focused mainly on class A sortases based on the assumptions that enzymes from other bacteria might reveal altered substrate

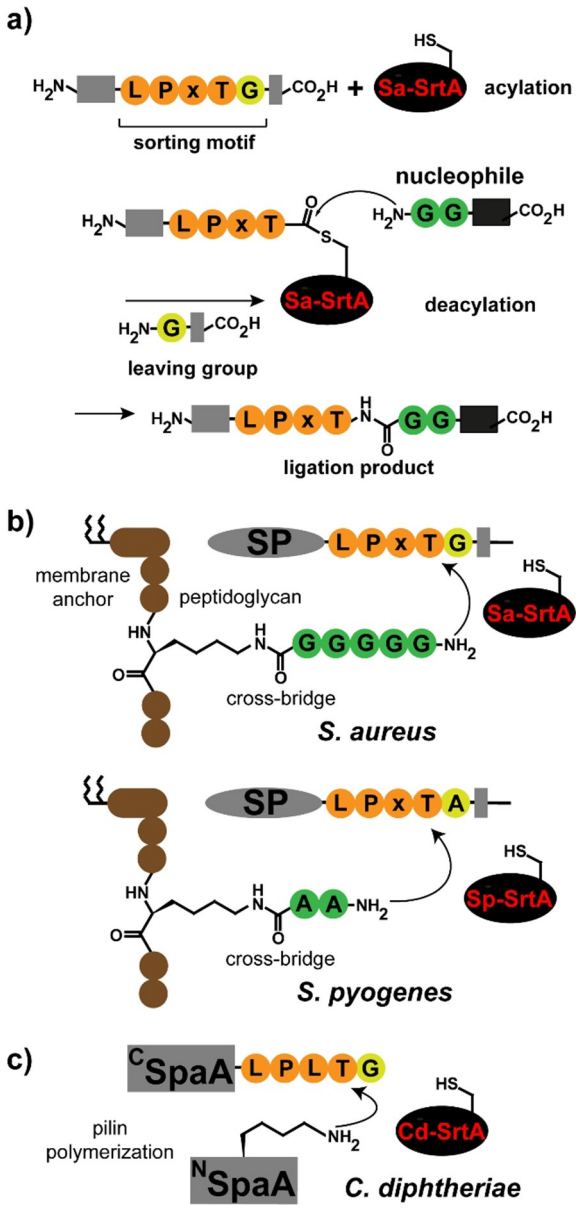

Figure 1. Sortases and sortase-mediated ligation (SML). a) Sortase A of *S. aureus* (Sa-SrtA) recognizes LPxTG sorting motifs as substrates and cleaves this sequence at the threonine residue with formation of a sortase-bound thioester. The downstream peptide with N-terminal glycine residue is released from the enzyme. The thioester is resolved by ligation to a peptide nucleophile with at least two glycine residues at the N terminus. b) Native sortases install surface proteins to the bacterial peptidoglycan cross-bridge. In the case of *S. aureus*, the cross-bridge contains a pentaglycine moiety as the nucleophile, while the peptidoglycan of *S. pyogenes* provides an N-terminal alanine residue for the ligation reaction. The composition of the bacterial peptidoglycan impacts sortase substrate specificity. c) Sortase A of *Corynebacterium diphtheriae* catalyzes isopeptide formation at an internal lysine of SpaA with a C-terminal sorting motif from a neighboring SpaA during pilin polylmerization. SP: surface protein.

specificity. The pentaglycine unit at the peptidoglycan cross-bridge deacylates Sa-SrtA bound thioesters in *S. aureus*, but this peptide nucleophile varies with the compositions of peptidoglycan.[2,5,11] For example, the peptidoglycan cross-bridge of *Streptococcus pyogenes* harbors a terminal alanine, serine, or glycine residue.[12] The corresponding *S. pyogenes* sortase A (Sp-SrtA) was shown to ligate the enzyme-bound thioesters to N-Ala nucleophiles (Figure 1b). This distinct specificity of Sp-SrtA could be exploited for a dual ligation scheme when combined with Sa-SrtA.[13]

Recent investigations aimed to explore the potential of previously uncharacterized class A sortases from various bacteria on a larger scale.[14] Focusing on the enzymes of *Staphylococcus* and *Streptococcus* origin, several sortase A genes were cloned, expressed, and the enzymes characterized with libraries of putative substrate peptides. These investigations uncovered that several *Streptococcus* sortases display a much more promiscuous substrate specificity than the Sa-SrtA. The experiments confirmed that *Streptococcus* sortases accept a broader range of substrates including LPxTG, LPxTA, and LPxTS motifs and N-Gly, N-Ala, and N-Ser nucleophiles.[14] In addition, alternative transacylation sites have been detected for some of these enzymes.[14a] Furthermore, the majority of investigated *Streptococcus* enzymes show a strong preference for an alternative LPxLG sorting motif, which is ligated more efficiently than the canonical LPxTG sorting motif *in vitro*.[14b]

A further promising enzyme for protein chemistry is sortase A of *C. diphtheriae* (Cd-SrtA). This enzyme is responsible for pilin polymerization in this bacterium. The deacylating amine nucleophile in Cd-SrtA catalyzed ligations is not the N-terminal residue of a peptidoglycan cross-bridge, but an internal lysine residue in pilin subunit SpaA. Pilin polymerization occurs by Cd-SrtA catalyzed crosslinking of a C-terminal LPLTG sequence in ^{C}SpaA to the internal lysine moiety in a neighboring ^{N}SpaA protein (Figure 1c). *In vitro* transpeptidase activity of Cd-SrtA could be established by introducing two mutations into an N-terminal lid extension of the enzyme that otherwise blocks the active site.[15] Introduction of three further mutations resulted in the Cd-SrtA-M3 enzyme that requires the internal lysine of SpaA as the ligation site and peptides with C-terminal LPLTG extension.[16] Protein bioconjugation with Cd-SrtA-M3 was then established by fusing SpaA to a protein of interest, followed by SML with synthetic LPLTG containing peptides. Importantly, Cd-SrtA-M3 and Sa-SrtA ligation schemes can be combined for dual labeling strategies when the protein of interest is fused to SpaA and contains an additional N-terminal glycine extension.

These investigations demonstrate that natural sortases possess new and useful features for SML. Given the number of sortase A sequences deposited in gene repositories, it appears likely that further investigative endeavors into the pool of uncharacterized sortases will uncover new biocatalysts with advantageous properties for SML (Table 1).

3. Engineering of Sortases

Sortase engineering is a highly promising strategy for establishing new biocatalysts optimized for the needs of SML approaches. Such tailor-made sortases with engineered features have been generated by directed protein evolution.

Table 1. Summary of sortases and sortase mutants discussed in this review.

	Preferred sorting motif	Preferred nucleophile	Activity compared to wt Sa-SrtA	Ca^{2+} dependent	Comments	Ref.
Wild-type sortases						
Sa-SrtA	LPxTG	N-Gly-Gly		yes		
Sp-SrtA	LPxLG	N-Ala-Ala	-	no	LPxTG, LPxTA donors and N-Gly-Gly, N-Ala-Gly, N-Ser-Gly acceptors are further substrates of Sp-SrtA	[14]
Cd-SrtA	LPLTG	Internal Lys of SpaA	n.d.	no		[16]
Sortase mutants with enhanced activity						
SrtA-M4	LPxTG	N-Gly-Gly	++	yes	k_{cat}/K_M (LPETG) = 28.000 $M^{-1}s^{-1}$ compared to 200 $M^{-1}s^{-1}$ for wt. 9.3-fold lower k_{cat}/K_M (GGG)	[19]
SrtA-M5	LPxTG	N-Gly-Gly	++	yes	k_{cat}/K_M (LPETG) = 23.000 $M^{-1}s^{-1}$ compared to 200 $M^{-1}s^{-1}$ for wt. 5.8-fold lower k_{cat}/K_M (GGG)	[19]
SrtA-M7	LPxTG	N-Gly-Gly	++	no	k_{cat}/K_M (LPETG) = 5.780 $M^{-1}s^{-1}$ in the absence of Ca^{2+}	[23]
SrtA-M5/D124G	LPxTG	N-Gly-Gly	++	yes	SML at antibody N termini; k_{cat}/K_M (LPETG) = 9.305 $M^{-1}s^{-1}$ compared to 159 $M^{-1}s^{-1}$ for wt.	[20]
SrtA-M5/D124G/Y187L/E189R	LPxTG	N-Gly-Gly	++	yes	SML at antibody C termini: k_{cat}/K_M (LPETG) = 16.722 $M^{-1}s^{-1}$.	[20]
SrtA-M4/R159N/K162P	LPxTG	N-Gly-Gly	++	yes	increased thermostability of M4	[37]
SrtA-P94H/A104T/E105D/G167E/Q172H	LPxTG	N-Gly-Gly	+	yes		[21]
SrtA-M3	LPxTG	N-Gly-Gly	+	yes	More DMSO resistent	[24]
Sortase mutants with altered specificity						
SrtA-F40	APxTG	N-Gly-Gly	−	yes		[25]
SrtA-A1-22	APxTG	N-Gly-Gly	−	yes		[26]
SrtA-F1-20	FPxTG	N-Gly-Gly	−	yes		[26]
eSrtA(4S)	LPxSG	N-Gly-Gly	++	yes	engineered on SrtA-M5 template	[27]
eSrtA(2A)	LAxTG	N-Gly-Gly	++	yes	engineered on SrtA-M5 template	[27]
mSrt2A	LAxTG	N-Gly-Gly	n.d.	no	engineered on eSrt(2A) template	[28]
Further sortase mutants						
Sa-SrtA-E105K/E108A	LPxTG	N-Gly-Gly	(−)	no		[29]
Sa-SrtA-E105K/E108Q	LPxTG	N-Gly-Gly	(−)	no		[29]

n.d. not determined in comparison to wild-type Sa-SrtA, ++ strongly enhanced, + enhanced, − reduced, (−) slightly reduced.

3.1. Engineered Sa-SrtA mutants with enhanced kinetics

Sa-SrtA possesses low catalytic activity as result of a high K_M of 5 mM for the LPxTG substrate and a turnover rate of 1 s^{-1}.[3] The high K_M in the millimolar range is a major obstacle for SML approaches and might be explained by the fact that native Sa-SrtA, which includes a membrane anchor, operates at high local concentrations of both substrates, which are also bound to the bacterial cell wall. Consequently, the evolutionary pressure on high substrate affinity might have been low. The lack of high local substrate concentrations in SML has been addressed by the development of proximity-based SML.[17] Sortase and LPxTG containing substrate proteins are tethered together by the SpyCatcher-SpyTag peptide-protein pair in this approach, resulting in accelerated ligation rates.[18]

However, protein evolution is a highly promising alternative approach for addressing this problem by enhancing the kinetics of the catalyst itself. In order to evolve more active Sa-SrtA mutants, an elaborated yeast display system for selecting sortase mutants with enhanced kinetics was established. Sortase mutants are displayed on the surface of yeast cells as fusion proteins with mating factor Aga2p, while the genetic information is stored in the cell. The Aga2p-mut-SrtA fusion is disulfide-linked to a second protein, Aga1p, which can be equipped with a peptide nucleophile by bioconjugation employing a glycine functionalized S6 tag. Conjugation of the S6 tag requires the coenzyme A (CoA) triglycine substrate and 4′-phosphopantethein transferase Sfp as catalyst. Upon installation of a triglycine nucleophile at Aga1p, bound SrtA was shown to ligate biotin-LPxTG peptides to the triglycine

nucleophile in a pseudo-intramolecular reaction, thereby covalently linking the biotin handle to the yeast surface (Figure 2a).[19] This yeast display system was used to screen a library of 7.8×10^7 sortase mutants generated by PCR amplification with mutagenic dNTPs. The screen was aimed primarily at improving the poor affinity for the sorting motif by adding the biotin-LPxTG in decreasing concentrations. Upon attachment of a streptavidin-conjugated fluorophore to the biotin handle, the yeast cells were subjected to FACS sorting. After several rounds of selection, SrtA-M4 and SrtA-M5, enzymes carrying 4 and 5 mutations, respectively, were isolated and shown to have a more than 100-fold higher catalytic efficiency (k_{cat}/K_M LPETG) than wild-type Sa-SrtA (Table 1). The mutations of SrtA–M5 are located close to the binding groove of the substrate containing the sorting motif (P94R/D160N/D165A/K196T) and the β7–β8 loop (K190E) interacting with the nucleophile (Figure 2d).[19] The SrtA-M5 mutant is intensively used in SML approaches and further protein evolution approaches on the SrtA-M5 scaffold have been reported.[1a]

A different screening approach was able to evolve further Sa-SrtA mutants with enhanced kinetics.[20] At first, the wild-type Sa-SrtA gene was subjected to error prone PCR. In addition, saturation mutagenesis of the SrtA-M5 template was used for generating a second sortase library. Both libraries were

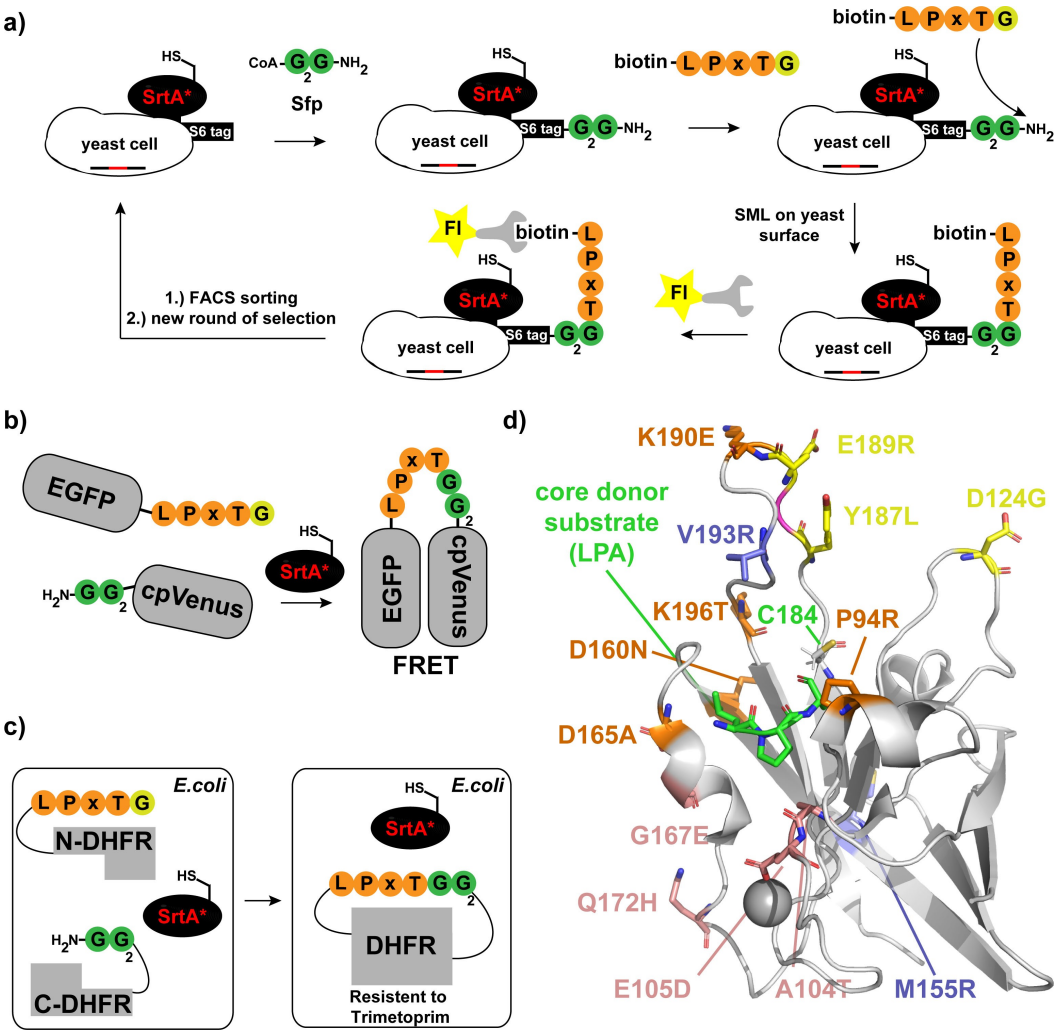

Figure 2. Selection schemes for sortase A activity enhancement. a) General scheme for evolving bond-forming enzymes by yeast display as applied to sortase A. SrtA of *S. aureus* is displayed as a fusion with the surface protein Aga2p (not shown), which itself is covalently linked to Agap1 (also not shown) carrying the reactive S6 peptide. Sfp phosphopantetheinyl transferase from *Bacillus subtilis* is then used to covalently link two glycine residues to the S6 tag, which provide the nucleophile for the sortase reaction. Biotinylated sorting motif peptide is then ligated to the N-terminal glycine residue by the spatially proximal sortase A protein, thereby coupling genotype with phenotype. Streptavidin-phycoerythrin is then used to detect and select successfully ligated product by FACS. b) a FRET-based platform for srtA activity screening. The LPETG sorting motif was fused to the C terminus of EGFP, while a triglycine moiety was attached to the N terminus of cpVenus. Bacterial clones expressing sortase variants leading to increased FRET signal intensity were selected. c) Protein complementation assay used for the directed evolution of sortase A. The two fragments of murine dihydrofolate reductase require covalent ligation by SrtA to confer activity and bacterial survival in the presence of the antibiotic trimethoprim. d) Activity enhancing mutations of SrtA from *S. aureus* depicted within the solution structure of the enzyme (PDB ID: 2KID). Mutations found in Chen et al. (2011)[19] are displayed in orange, those found in Chen et al. (2016)[20] in yellow, those from Suliman et al. (2017) [21] in pink and those from Wojcik et al. (2019)[22] in blue. Amino acids are labeled by type and number with respect to the *S. aureus* wt protein, while the corresponding mutations found in the respective screens are indicated behind the number. Residues 59–72 from the structure have been omitted for clarity. Cys184 indicates the active site, while the core LPA tripeptide from the peptide analog is shown in green.

screened for high FRET signal intensities resulting from SML of EGFP and cpVenus equipped with LPETG motif and triglycine moiety, respectively (Figure 2b). These screens uncovered three further mutations that enhanced sortase kinetics when combined with the M5 mutations (M5/D124G and M5/Y187L/E189R) (Figure 2d). These mutants were particularly useful for SML of antibodies with SrtA-M5/D124G as the most efficient mutant for ligating peptides to the N terminus of an antibody. SrtA-M5/Y187E/E189R was highly effective in ligation reactions at the C termini of antibodies.[20]

A further directed sortase evolution approach was based on dihydrofolate reductase (DHFR) complementation screens. This screening platform capitalizes on the assembly of full-length mammalian DHFR from two fragments extended with LPxTG and triglycine sequences by SML (Figure 2c). Upon inhibition of endogenous bacterial DHFR by Trimethoprim, the ligated full-length murine DHFR becomes essential for bacterial survival.

Screening a library of 2×10^7 sortase mutants generated by error-prone PCR uncovered five recurrent mutations P94H/A104T/E105D/G167E/Q172H) localized in the β6–β7 and β7–β8 loops of Sa-SrtA (Figure 2d).[21] P94 is the only common residue found between two of the screens, supporting the conclusion that the type and exact conditions of the screening and selection methods influence the results. More interestingly, almost all mutations in these different screens change the charge of the corresponding wild-type amino acid, indicating that additional long-ranging electrostatic interactions play a pivotal role in substrate recognition and catalysis. Thus, while several mutations localize in the vicinity of the substrates, others are found more remotely at the tip of the β7–β8 loop (Figure 2d).

Finally, further screening platforms for directed sortase evolution are currently in development. A recently reported strategy makes use of immobilized reporters of SML reactions in microtiter plate formats.[30] SML is used to ligate a reporter protein to peptides, serving as handles for immobilization. The reporter protein catalyzes a colorimetric readout reaction when successfully immobilized in the plates. Screening sortase libraries obtained by saturation mutagenesis at selected sites confirmed the efficiency of the M5 mutations for Sa-SrtA catalysis (Table 1).[30]

3.2. Sortase mutants with altered substrate specificity

In addition to engineering Sa-SrtA kinetics, this enzyme has been also subjected to directed evolution of substrate specificity.

Sortases with altered substrate specificity are attractive for protein semisynthesis approaches where proteins are reconstituted form synthetic and recombinant fragments by ligation reactions. Sortases with altered specificity could reduce the number of mutations that remain as 'ligation scarring' with the ligation products. Furthermore, sortases with alternative sorting motifs and nucleophile specificities could enable orthogonal SML with more than two fragments.

Initial attempts to change the Sa-SrtA specificity were based on the observation that Sa-SrtB recognizes an NPQTN sorting motif but possesses very limited *in vitro* activity. Structural data allowed residues of the β6–β7 loop to be assigned as the main contact site between sortase and the substrate containing the sorting motif. A loop-swap chimera (SrtLS) was generated by grafting the β6–β7 loop of Sa-SrtB onto the Sa-SrtA protein scaffold.[31] Chimeric SrtLS recognized the NPQTN sequence as substrate, but was only able to catalyze cleavage of this motif and not the ligation reaction. A more recently developed Sp-SrtA chimera was engineered, focusing on altered specificity for the sortase deacylating peptide nucleophile.[22] Nucleophile recognition is largely mediated by residues of the β7–β8 loop. By grafting this region from Sa-SrtA onto Sp-SrtA, the resulting hybrid enzyme is restricted to N-Gly substrates. The hybrid enzyme was no longer able to ligate LPxTA motifs to N-Ala nucleophiles like wild-type Sp-SrtA, confirming the important role of the β7–β8 loop in nucleophile substrate recognition.

The first directed evolution of Sa-SrtA substrate specificity was used in a phage display system for selecting mutants with altered sorting motif specificity. A library of 1×10^8 Sa-SrtA mutants was generated by NNK mutagenesis, randomizing six residues of the β6–β7 loop predicted to interact with the leucine residue of the LPxTG motif (Figure 3a).[25] The library was displayed as C-terminal fusion with the pIII surface protein on M13 phages. The N terminus of the displayed sortase mutants was extended by a pentaglycine tail, enabling ligations to the N termini. Ligation of biotinylated LPxTG peptides allowed enrichment of phages on streptavidin resins. After three rounds of phage display with a biotin-FPxTG bait, isolated sortase mutants were characterized. In general, these mutants were promiscuous with respect to the leucine position in the sorting motif and ligated the expected FPxTG sequence. The activity of the selected mutants was generally lower than that of the wild-type enzyme. One mutant, referred to as F40 sortase, appeared particularly interesting because this enzyme catalyzed ligation reactions of APxTG and DPxTG substrates. The first of these amino acid sequences is present in histone H3, enabling traceless semisynthesis of this protein by SML.[25] However, the general broadening of specificity in this phage display campaign was surprising and further investigated. The initial library design was based on the first reported Sa-SrtA crystal structure.[32] However, the subsequently elucidated NMR structure of the sortase A construct with covalently attached LPA core motif analog revealed a very different conformation of the β6–β7 loop in the substrate bound state, indicating that the residues selected for randomization might not have been ideal for evolving the Sa-SrtA specificity.[33] This NMR structure could retrospectively rationalize the preference of Ala and Asp instead of Leu at the first position of the sorting motif, despite the ligation reaction having been performed with an FPKTG peptide. A more bulky Met residue is found instead of Val at position 168 of the β6–β7 loop in the F40 sortase, likely to geometrically favor small side-chains in the substrate. In addition, R197 of the β7–β8 loop likely provides a favorable charge interaction in case the N-terminal residue of the sorting motif is an Asp. To more rationally change the properties of the

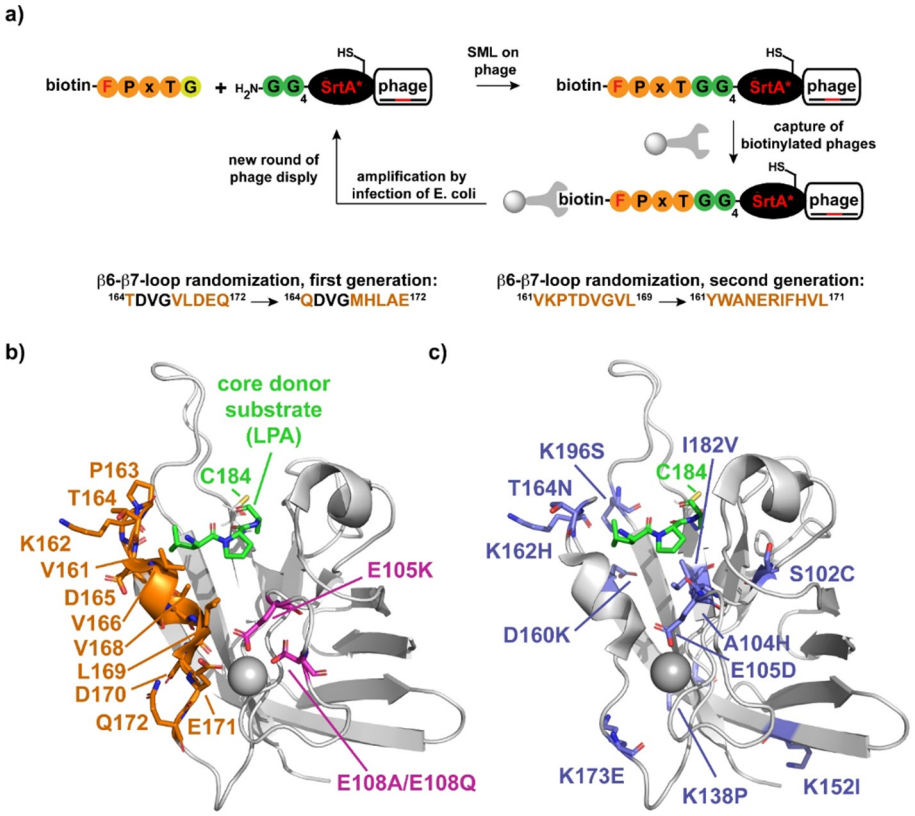

Figure 3. Sortase mutants with altered substrate specificity. a) Selection scheme for changing the specificity of SrtA. Self-ligation of an altered sorting motif to the N-terminal glycine artificially fused to the randomized sortases displayed on phage allows active sortases to be labeled by biotin and subsequent enrichment on NeutrAvidin-coated agarose beads. The design of the first-generation SrtA library is shown below left and indicates the six NNK randomized positions in orange. The second-generation library was randomized at positions 161–169 by using trinucleotides and excluding stop and cysteine codons. Furthermore, loop lengths were allowed to vary between 7 and 11 amino acids. The most potent variant indeed contained 11 amino acids, as indicated bottom right. b) Altering sortase A specificity and calcium dependence by design. The amino acids of SrtA that were randomized or changed by design are rendered as sticks. Libraries selected for the recognition of amino acids other than leucine at the first position of the sorting motif are shown in orange with the randomized region restricted to the β6–β7 loop, see (a) for the identified mutations. Amino acids are indicated by type and number and the core LPA motif of the ligand is displayed in green. The two mutations that were rationally designed to render SrtA Ca^{2+}-independent are shown in purple. The bound calcium is displayed as a gray sphere. c) Mutations identified by full randomization of SrtA and selected to ligate the sorting motif LAETG are shown in blue. As in (b) amino acids that were found to be mutated compared to wt sortase A from *S. aureus* are rendered as sticks (in blue). In the screen performed in this study,[20] randomization started from sortase A with enhanced kinetic properties (the so-called M5 variant or eSrtA, Figure 2d), thus two of the mutations, N160K and T196S, affected residues at the M5 variant positions.

β6–β7 loop, a re-designed Sa-SrtA library was generated, based on the NMR Sa-SrtA structure using trinucleotide building blocks, and also including variations of the β6–β7 loop length.[26] Phage display screening of this re-designed library with FPxTG and APxTG baits yielded sortase mutants SrtA-F1-20 and SrtA-A1-22 with expected preferences for these altered sorting motifs. The majority of isolated mutants contained β6–β7 loops that were longer than the native loop (11 instead of nine residues; Figure 3b).[26]

A set of orthogonal sortases recognizing LAxTG and LPxSG motifs evolved on the basis of the yeast display system discussed above (Figure 2a) was reported, enabling dual labeling of proteins with SML. The screening campaign of a sortase library at 2% mutation level was optimized for high selectivity, and yielded sortase mutants with the desired altered specificity and no significant loss in activity.[27] Three key mutations responsible for substrate promiscuity were derived from this screen and used to generate sortase mutants accepting LAxTG and LPxSG as sorting motifs. These mutants were subjected to further directed evolutions, finally delivering evolved sortases eSrtA(2A) and eSrtA(4S) showing high specificity for LAxTG and LPxSG motifs. Both enzymes accumulated a significant number of mutations during the evolution process, some of which were located at contact sites between the enzyme and the substrate, while other mutations were found in remote regions (Figure 3c). Importantly, these mutants maintained a high increase in activity (Table 1).[27]

3.3. Calcium-independent sortase mutants and mutants with improved solvent or temperature stability

A further challenge of Sa-SrtA is the requirement of calcium for effective catalysis. Calcium-dependency constitutes a significant problem for applications in living cells where calcium levels are low. Calcium-dependency is observed mainly for sortases of

Staphylococci origin, while other sortases, including Sp-SrtA, are calcium-independent. However, since Sa-SrtA is the most active sortase uncovered so far, a calcium-independent enzyme was generated by rational design.[29] In Sa-SrtA, the Ca^{2+} ion is important for stabilizing the closed conformation of the β6–β7 loop in a substrate-bound state. The Ca^{2+} binding site is formed by three glutamic acid residues: Glu171 of the β6–β7 loop, and Glu 108 and Glu 105 located in the β3–β4 loop. In the calcium-independent Sp-SrtA, the Ca^{2+} interactions are substituted by a salt bridge between a lysine and an aspartic acid moiety. The calcium-independent Sa-SrtA was generated by installing a similar salt bridge into the β3–β4 loop. The resulting mutants, Sa-SrtA(E105K/E108A) and Sa-SrtA(E105K/E108Q), catalyzed SML reactions with slightly slower kinetics compared to the wild-type enzyme in the absence of calcium. Mutations mediating calcium independence could be further combined with the rate-accelerating mutations of the SrtM5 enzyme.[34] The resulting SrtA-M7 combines the advantageous features of both engineering approaches and is now a widely used tool in protein chemistry.[23] The SrtA-M7 enzyme was further shown to accept a broad range of small molecule amine nucleophiles containing biorthogonal groups. When co-expressing SrtA-M7 and LPxTG containing proteins in *Escherichia coli* this feature was exploited for the in vivo incorporation of alkynes, azides and tetrazines into the co-expressed proteins.[35]

A further desirable feature of enzymes used for biotechnology is their resistance to organic solvents. SML in organic solvents or mixtures of water and organic solvents would allow ligation reactions with hydrophobic substrates. Wild-type sortase activity decreases dramatically in the presence of high concentrations of solvents like DMSO. An engineering campaign screened for sortase mutants able to perform SML in the presence of 45% DMSO. The screen was based on ligation of a LPxTG modified CueO laccase reporter to an N-Gly modified anchoring protein, enabling immobilization of the ligation products in polypropylene microtiter plates. A sortase library generated by saturation mutagenesis was screened in the presence of 45% DMSO for efficient SML. The screen delivered sortase mutants, including SrtA-M3, which showed higher transpeptidase activity than wild type Sa-SrtA in both the presence and absence of 45% DMSO.[24]

To improve the thermostability of Sa-SrtA, multiple sequence alignments of sortases revealed certain positions of sortases from extremophiles to be conserved or distinct from Sa-SrtA, identifying M155V and V193R as mutations that increased activity at 37 °C or higher (Figure 2d; Table 1).[36]

Notably, sortases evolved for high catalytic activity commonly come at the cost of decreased thermostability. It was therefore attempted to improve the stability of the activity-enhanced M4 mutant from the original yeast display screen (Table 1).[19,37] This mutant shows a more than 10 °C lower thermal melting temperature compared to wild-type SaSrtA (48.6 vs 59.4 °C). Single site-saturation mutagenesis of residues in the β6–β7 loop identified two mutations (R159N and K162P) that, when combined, led to a 6 °C higher stability of this new M6-SrtA. The relative activity of the M6 mutant was increased 58-fold compared to wild-type SaSrtA after pre-incubation of the enzymes at 55 °C for 1 hour. Further improvement of stability by 1.5 °C was achieved by head-to-tail cyclization of the engineered variant.[37]

4. Applications of Engineered Sortases

Engineered sortases have already made a strong impact on protein chemistry. In particular, SrtA-M5 and the calcium-independent SrtA-M7 with enhanced kinetics are often used as more efficient alternatives to wild-type Sa-SrtA. Biotechnological approaches, such as the generation of antibody conjugates with biophysical probes or drugs by SML, benefit strongly from the accelerated ligation reactions catalyzed by these enzymes.[1a,f]

Here we will focus on biochemical applications of engineered sortases, particularly the utility of these enzymes for investigating the function of post-translational modifications (PTMs) of proteins. We will refer to other reviews for a more global view on sortase applications.[1]

PTMs can serve as regulators of protein function and activity by modulating protein-protein interactions or inducing conformational changes within the protein structure.[38] Histone proteins package eukaryotic DNA into chromatin and are extensively decorated by PTMs in the N-terminal tail regions.[39] These histone tails protrude from the basic structural units of chromatin, the nucleosomes, into the nucleus.[40] Modifications of histone tails constitute an epigenetic histone code by recruiting chromatin factors to the DNA template. Protein semisynthesis is an attractive tool for generating histones with defined modification patterns. In the first step, PTMs are introduced into the synthetic fragments and a subsequent ligation reaction introduces the modification into the full-length protein.[39] SML is particularly useful for generating semisynthetic histone H3. Histone H3 contains the sequence APATG at the interface between the tail and the globular fold, spanning positions 29 to 33, which resembles the LPxTG sorting motif closely. Wild-type sortase has been used for generating libraries of nucleosomes with defined methylation and phosphorylation patterns.[41] The H3 tails were installed on the level of assembled nucleosomes and allowed the generation of libraries with defined modification patterns. Library screens with chromatin reader protein HP1 uncovered a cross-talk between the HP1 recruitment site at H3K9me2 and phosphorylation of the remote residue S28, thereby demonstrating the utility of SML for chromatin research.[41]

However, SML of H3 with wild-type Sa-SrtA comes at the price of an A29L point mutation introduced into H3, which could potentially impact the native interactions between chromatin factors and the histone tail. The F40 sortase mutant represents an alternative catalyst for H3 SML, despite being less active than wild-type Sa-SrtA. Ligation of the native H3 sequence APATG by F40-SrtA allows traceless H3 semisynthesis and has already been applied in chromatin biochemistry.[25] The SAGA complex is a transcriptional coactivator and possesses lysine acetyltransferase activity.[42] It was shown that SAGA is recruited to nucleosomes containing trimethylated H3K4me3

by a tandem Tudor domain with SAGA protein Sgf29. H3K4me3 nucleosomes were generated by F40-SrtA catalyzed SML of H3 from a synthetic H3K4me3 peptide and N-terminal truncated recombinant H3.[43] In contrast to the library approach discussed above, this SML reaction was performed on the histone level, and the ligated H3 was subsequently incorporated into reconstituted nucleosomes. In a competitive setting of H3K4me3 nucleosomes in the presence of unmodified nucleosomes, SAGA preferentially hyperacetylated the H3K4me3 fraction, allowing a model to be established for the recruitment of the SAGA complex to chromosomal promoter regions.[43]

In addition to lysine acetyltransferases, their counterparts, lysine or histone deacetylases (HDACs) constitute a further important class of transcriptional regulators. The CoREST complex plays a central role in neuronal development and possesses deacetylase activity, mediated by HDAC1. In addition, the complex contains LSD1, a histone demethylase of H3K4me2.[44] HDAC1 and LSD1 are tethered by scaffolding protein CoREST, forming the complex core referred to as LHC. The catalytic properties and regulation of LHC were characterized using nucleosomes modified at H3 with methylation marks at K4 and further acetylated lysine residues at the K9, K14, and K18 positions.[45] Modified H3 was generated by SML with F40 sortase and subsequently incorporated into nucleosomes (Figure4 a). The SML reaction involved sorting motif substrates with a depsipeptide bond downstream of the threonine residue in the sorting motif. The released leaving group contains an N-terminal glycolic acid moiety, which does not compete with the N-Gly nucleophile of truncated H3, thus driving the SML reaction to completion.[8a] This strategy was also applied for establishing nucleosomal libraries with wild type Sa-SrtA as discussed above. The site-specifically methylated and acetylated nucleosomes showed that H3K14ac suppresses demethylation of H3K4me2 by LSD1. Furthermore, HDAC1 removes the H3K14ac mark sluggishly compared to acetylation marks at H3K9ac and H3K18ac. These observations indicate a regulatory crosstalk between H3K4me2 and H3K14ac.[45] The SML strategy was further applied for generating H3-acetylated nucleosomes for investigating further complexes containing class I HDACs including CoREST, NuRD, Sin3B, MiDAC, and SMART.[46] Deacetylation experiments showed variations in substrate specificities of these complexes and a glycine residue upstream of H3K14ac that induces reduced deacetylase activity of CoREST towards this acetylation mark.[47]

Ubiquitylation and SUMOylation are two further important PTMs of proteins.[48] These small proteins are linked to lysine side chains via isopeptide bonds. Several elaborate ligation strategies have been developed for site-specific ubiquitylation and SUMOylation of proteins.[49] A recently reported strategy, referred to as "sortylation", combines genetic-code expansion and SML with engineered sortases enabling ubiquitylation and SUMOylation in living cells (Figure 4b).[28] This approach was made possible by genetic encoding of lysine carrying an azido-Gly-Gly isopeptide extension using amber suppression technology. Upon incorporation into recombinant proteins, the azide can be reduced to an amine to serve as a nucleophile in SML reactions. To this end, the LPxTG motif was engineered into the C terminus of ubiquitin or SUMO, allowing their ligation with SrtA-M5. However, the inserted proline residue yielded an undesirable outcome and was substituted with an alanine residue, resulting in an LAxTG motif serving as substrate of engineered eSrtA(2A). The resulting ubiquitylated proteins differ at two positions from the native amino acid sequence, and the isopeptide bond was shown to be resistant to deubiquitinases. This methodology could be further established in living cells by coexpression of all reaction partners. However, in order to cope with the low intracellular calcium concentration, the eSrtA(2A) enzyme was subjected to a further engineering step, introducing the D105H and E108E mutations to render the resulting mSrt2A calcium-independent. The mSrt2A enzyme allowed

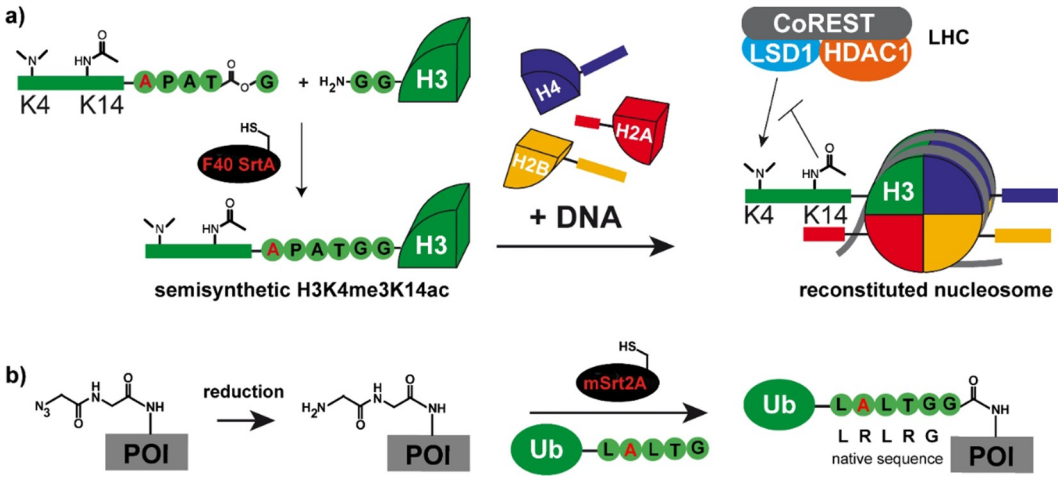

Figure 4. Applications of engineered sortases. a) Engineered F40 sortase allows traceless semi-synthesis of histone H3. Highly efficient SML of H3 was enabled by inserting a depsipeptide bond into the sorting motif. Designer nucleosomes site-specifically methylated and acetylated at Lys4 and Lys14 allowed mechanistic insights into the recognition of the LHC core of the CoREST complex. b) Genetic encoding of a lysine residue with isopeptide azido-Gly-Gly modification allowed in vivo installation of ubiquitin by engineered mSrt2A. Reduction of the azido moiety delivered the amine nucleophile, which was ligated with ubiquitin containing a C-terminal LALTG sequence, which differed only at two positions from the native sequence.

application of this methodology in live cells, thus bypassing the native ubiquitylation and SUMOylation systems.[28]

These recent examples of SML with engineered sortases provide a first glimpse at the potential of these tools for investigating protein modifications. With more engineered sortases providing new features, it appears likely that the impact of SML on modern biochemistry will continue to evolve.

5. Summary and Outlook

Sortase engineering has come of age, showing that wild-type sortase can be improved with regard to its catalytic activity and that its specificity can be changed with regard to individual positions in the sorting motif. This allows for interesting applications in the field of intracellular signaling, probing the role of individual post-translational modification events. Capitalizing on the already established set of engineered sortases allows researchers to create sortases with new properties by combining structurally independent elements of the enzyme (e.g., the incorporation of kinetically derived mutants into specificity changing mutants). Alternatively, future research may exploit the richness of natural sortases by methods such as DNA shuffling, with the promise to further explore sequence space and to direct evolution towards new substrate specificities. These attempts are ideally flanked by mechanistic investigations that capture the essential dynamics and structural transitions that accompany the two ligation steps. Based on its ease of expression and manipulation, sortase A might become a good example of how to stretch the limits of a naturally occurring enzyme for the purpose of biotechnology.

Acknowledgements

We are grateful for support by priority program SPP 1623 of the Deutsche Forschungsgemeinschaft with grants FR-1325/15-1 (CF) and SCHW 1163/7-1 (DS). We further like to thank Katherine Maffucci for critical reading of the manuscript. Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: protein bioconjugation · protein engineering · protein semisynthesis · sortases · transpeptidases

[1] a) J. M. Antos, M. C. Truttmann, H. L. Ploegh, *Curr. Opin. Struct. Biol.* **2016**, *38*, 111–118; b) A. W. Jacobitz, M. D. Kattke, J. Wereszczynski, R. T. Clubb, *Adv. Protein Chem. Struct. Biol.* **2017**, *109*, 223–264; c) N. Pishesha, J. R. Ingram, H. L. Ploegh, *Annu. Rev. Cell Dev. Biol.* **2018**, *34*, 163–188; d) M. W. L. Popp, H. L. Ploegh, *Angew. Chem. Int. Ed.* **2011**, *50*, 5024–5032; *Angew. Chem.* **2011**, *123*, 5128–5137; e) M. Ritzefeld, *Chem. Eur. J.* **2014**, *20*, 8516–8529; f) L. Schmohl, D. Schwarzer, *Curr. Opin. Chem. Biol.* **2014**, *22*, 122–128.
[2] O. Schneewind, A. Fowler, K. F. Faull, *Science* **1995**, *268*, 103–106.
[3] R. G. Kruger, B. Otvos, B. A. Frankel, M. Bentley, P. Dostal, D. G. McCafferty, *Biochemistry* **2004**, *43*, 1541–1551.
[4] M. L. Bentley, E. C. Lamb, D. G. McCafferty, *J. Biol. Chem.* **2008**, *283*, 14762–14771.
[5] H. Ton-That, G. Liu, S. K. Mazmanian, K. F. Faull, O. Schneewind, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12424–12429.
[6] H. Y. Mao, S. A. Hart, A. Schink, B. A. Pollok, *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671.
[7] M. W. Popp, J. M. Antos, G. M. Grotenbreg, E. Spooner, H. L. Ploegh, *Nat. Chem. Biol.* **2007**, *3*, 707–708.
[8] a) D. J. Williamson, M. A. Fascione, M. E. Webb, W. B. Turnbull, *Angew. Chem. Int. Ed.* **2012**, *51*, 9377–9380; *Angew. Chem.* **2012**, *124*, 9511–9514; b) F. Liu, E. Y. Luo, D. B. Flora, A. R. Mezo, *J. Org. Chem.* **2014**, *79*, 487–492; c) Y. Yamamura, H. Hirakawa, S. Yamaguchi, T. Nagamune, *Chem. Commun. (Camb.)* **2011**, *47*, 4742–4744; d) R. David Row, T. J. Roark, M. C. Philip, L. L. Perkins, J. M. Antos, *Chem. Commun. (Camb.)* **2015**, *51*, 12548–12551; e) S. A. Reed, D. A. Brzovic, S. S. Takasaki, K. V. Boyko, J. M. Antos, *Bioconjugate Chem.* **2020**, *31*, 1463–1473.
[9] T. Spirig, E. M. Weiner, R. T. Clubb, *Mol. Microbiol.* **2011**, *82*, 1044–1059.
[10] S. K. Mazmanian, H. Ton-That, K. Su, O. Schneewind, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2293–2298.
[11] O. Schneewind, D. Missiakas, *Biophys. Biochim. Acta Mol. Cell Res.* **2014**, *1843*, 1687–1697.
[12] K. H. Schleifer, O. Kandler, *Bacteriol. Rev.* **1972**, *36*, 407–477.
[13] J. M. Antos, G. L. Chew, C. P. Guimaraes, N. C. Yoder, G. M. Grotenbreg, M. W. L. Popp, H. L. Ploegh, *J. Am. Chem. Soc.* **2009**, *131*, 10800–10801.
[14] a) K. D. Nikghalb, N. M. Horvath, J. L. Prelesnik, O. G. B. Banks, P. A. Filipov, R. D. Row, T. J. Roark, J. M. Antos, *ChemBioChem* **2018**, *19*, 185–195; b) L. Schmohl, J. Bierlmeier, N. von Kugelgen, L. Kurz, P. Reis, F. Barthels, P. Mach, M. Schutkowski, C. Freund, D. Schwarzer, *Bioorg. Med. Chem.* **2017**, *25*, 5002–5007.
[15] C. Y. Chang, B. R. Amer, J. Osipiuk, S. A. McConnell, I. H. Huang, V. Hsieh, J. Fu, H. H. Nguyen, J. Muroski, E. Flores, R. R. O. Loo, J. A. Loo, J. A. Putkey, A. Joachimiak, A. Das, R. T. Clubb, H. Ton-That, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E5477–E5486.
[16] S. A. McConnell, B. R. Amer, J. Muroski, J. Fu, C. Chang, R. R. Ogorzalek Loo, J. A. Loo, J. Osipiuk, H. Ton-That, R. T. Clubb, *J. Am. Chem. Soc.* **2018**, *140*, 8420–8423.
[17] H. H. Wang, B. Altun, K. Nwe, A. Tsourkas, *Angew. Chem. Int. Ed.* **2017**, *56*, 5349–5352; *Angew. Chem.* **2017**, *129*, 5433–5436.
[18] H. H. Wang, A. Tsourkas, *Methods Mol. Biol.* **2019**, *2008*, 165–177.
[19] I. Chen, B. M. Dorr, D. R. Liu, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11399–11404.
[20] L. Chen, J. Cohen, X. Song, A. Zhao, Z. Ye, C. J. Feulner, P. Doonan, W. Somers, L. Lin, P. R. Chen, *Sci. Rep.* **2016**, *6*, 31899.
[21] M. Suliman, V. Santosh, T. C. M. Seegar, A. C. Dalton, K. M. Schultz, C. S. Klug, W. A. Barton, *PLoS One* **2017**, *12*, e0184271.
[22] M. Wojcik, K. Szala, R. van Merkerk, W. J. Quax, Y. L. Boersma, *Proteins* **2020**, *88*, 1394–1400.
[23] H. Hirakawa, S. Ishikawa, T. Nagamune, *Biotechnol. J.* **2015**, *10*, 1487–1492.
[24] Z. Zou, H. Alibiglou, D. M. Mate, M. D. Davari, F. Jakob, U. Schwaneberg, *Chem. Commun. (Camb.)* **2018**, *54*, 11467–11470.
[25] K. Piotukh, B. Geltinger, N. Heinrich, F. Gerth, M. Beyermann, C. Freund, D. Schwarzer, *J. Am. Chem. Soc.* **2011**, *133*, 17536–17539.
[26] L. Schmohl, J. Bierlmeier, F. Gerth, C. Freund, D. Schwarzer, *J. Pept. Sci.* **2017**, *23*, 631–635.
[27] B. M. Dorr, H. O. Ham, C. An, E. L. Chaikof, D. R. Liu, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13343–13348.
[28] M. Fottner, A. D. Brunner, V. Bittl, D. Horn-Ghetko, A. Jussupow, V. R. I. Kaila, A. Bremm, K. Lang, *Nat. Chem. Biol.* **2019**, *15*, 276–284.
[29] H. Hirakawa, S. Ishikawa, T. Nagamune, *Biotechnol. Bioeng.* **2012**, *109*, 2955–2961.
[30] Z. Zou, D. M. Mate, K. Rubsam, F. Jakob, U. Schwaneberg, *ACS Comb. Sci.* **2018**, *20*, 203–211.
[31] M. L. Bentley, H. Gaweska, J. M. Kielec, D. G. McCafferty, *J. Biol. Chem.* **2007**, *282*, 6571–6581.
[32] Y. N. Zong, T. W. Bice, H. Ton-That, O. Schneewind, S. V. L. Narayana, *J. Biol. Chem.* **2004**, *279*, 31383–31389.
[33] N. Suree, C. K. Liew, V. A. Villareal, W. Thieu, E. A. Fadeev, J. J. Clemens, M. E. Jung, R. T. Clubb, *J. Biol. Chem.* **2009**, *284*, 24465–24477.
[34] I. Wuethrich, J. G. Peeters, A. E. Blom, C. S. Theile, Z. Li, E. Spooner, H. L. Ploegh, C. P. Guimaraes, *PLoS One* **2014**, *9*, e109883.

[35] J. E. Glasgow, M. L. Salit, J. R. Cochran, *J. Am. Chem. Soc.* **2016**, *138*, 7496–7499.
[36] M. Wojcik, S. Vazquez Torres, W. J. Quax, Y. L. Boersma, *Protein Eng. Des. Sel.* **2019**, *32*, 555–564.
[37] Z. Zou, D. M. Mate, M. Noth, F. Jakob, U. Schwaneberg, *Chemistry* **2020**.
[38] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto Jr., *Angew. Chem. Int. Ed.* **2005**, *44*, 7342–7372; *Angew. Chem.* **2005**, *117*, 7508–7539.
[39] a) B. Fierz, M. G. Poirier, *Annu. Rev. Biophys.* **2019**, *48*, 321–345; b) W. Fischle, H. D. Mootz, D. Schwarzer, *Curr. Opin. Chem. Biol.* **2015**, *28*, 131–140; c) M. M. Muller, T. W. Muir, *Chem. Rev.* **2015**, *115*, 2296–2349.
[40] a) N. Justin, V. De Marco, R. Aasland, S. J. Gamblin, *Curr. Opin. Struct. Biol.* **2010**, *20*, 730–738; b) S. Rea, F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis, T. Jenuwein, *Nature* **2000**, *406*, 593–599; c) B. D. Strahl, C. D. Allis, *Nature* **2000**, *403*, 41–45; d) S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, D. J. Patel, *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025–1040.
[41] D. A. Pelaz, H. Mahler, D. Schwarzer, W. Fischle, *Protein Sci.* **2015**, *24*, 52–52.
[42] J. H. M. Soffers, V. V. Popova, J. L. W. Workman, *Trends Biochem. Sci.* **2020**, *45*, 547–549.
[43] A. E. Ringel, A. M. Cieniewicz, S. D. Taverna, C. Wolberger, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E5461–5470.
[44] a) S. Maksour, L. Ooi, M. Dottori, *eNeuro* **2020**, *7*, ENEURO.0337-19.2020; b) I. A. Qureshi, S. Gokhan, M. F. Mehler, *Cell Cycle* **2010**, *9*, 4477–4486.
[45] M. Wu, D. Hayward, J. H. Kalin, Y. Song, J. W. Schwabe, P. A. Cole, *eLife* **2018**, *7*, doi: 10.21775/cimb.037.001.
[46] M. Bantscheff, C. Hopf, M. M. Savitski, A. Dittmann, P. Grandi, A. M. Michon, J. Schlegl, Y. Abraham, I. Becher, G. Bergamini, M. Boesche, M. Delling, B. Dumpelfeld, D. Eberhard, C. Huthmacher, T. Mathieson, D. Poeckel, V. Reader, K. Strunk, G. Sweetman, U. Kruse, G. Neubauer, N. G. Ramsden, G. Drewes, *Nat. Biotechnol.* **2011**, *29*, 255-U124.
[47] Z. A. Wang, C. J. Millard, C. L. Lin, J. E. Gurnett, M. Wu, K. Lee, L. Fairall, J. W. Schwabe, P. A. Cole, *eLife* **2020**, *9*.
[48] D. Komander, M. Rape, *Annu. Rev. Biochem.* **2012**, *81*, 203–229.
[49] a) M. Jbara, H. Sun, G. Kamnesky, A. Brik, *Curr. Opin. Chem. Biol.* **2018**, *45*, 18–26; b) M. P. C. Mulder, K. F. Witting, H. Ovaa, *Curr. Issues Mol. Biol.* **2020**, *37*, 1–20.

Manuscript received: October 28, 2020
Revised manuscript received: December 7, 2020
Accepted manuscript online: December 8, 2020
Version of record online: February 3, 2021