JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Directed Evolution of Sortase A Mutants with Altered Substrate Selectivity Profiles

Kirill Piotukh,[†] Bernhard Geltinger,[‡] Nadja Heinrich,[∥] Fabian Gerth,[†] Michael Beyermann,[*,∥] Christian Freund,[*,†] and Dirk Schwarzer[*,‡]

Departments of [†]Protein Engineering, [‡]Protein Chemistry, and [∥]Peptide Synthesis, Leibniz-Institut für Molekulare Pharmakologie, Robert-Rössle-Strasse 10, 13125 Berlin, Germany

Supporting Information

ABSTRACT: The ligation of two polypeptides in a chemoselective manner by the bacterial transpeptidase sortase A has become a versatile tool for protein engineering approaches. When sortase-mediated ligation is used for protein semisynthesis, up to four mutations resulting from the strict requirement of the LPxTG sorting motif are introduced into the target protein. Here we report the directed evolution of a mutant sortase A that possesses broad substrate selectivity. A phage-display screen of a mutant sortase library that was randomized in the substrate recognition loop was used to isolate this mutant. The altered substrate selectivity represents a gain-of-function that was exploited for the traceless semisynthesis of histone H3. Our report is a decisive step toward a platform of engineered sortases with distinct ligation properties that will conceivably allow for more versatile assemblies of modified proteins in biotechnological approaches.

Chemoselective ligation techniques have endowed scientists with powerful tools for studying the structure−function relationship in proteins. These techniques enable site-specific modification, semisynthesis, and even the total chemical synthesis of proteins.[1] Complementary to chemical approaches like native chemical ligation, enzyme-mediated ligation schemes based on the bacterial transpeptidase sortase A have gained considerable attention.[2] Gram-positive bacteria use sortases to anchor surface proteins to their peptidoglycan layers.[3] The ligation process is reminiscent of that of cysteine protease reactions: In the first step, sortase A recognizes a highly conserved LPxTG sorting motif in the target protein and cleaves this sequence C-terminal to the Thr residue. The resulting thioester intermediate—instead of being hydrolyzed—is ligated to the penta-Gly unit of the bacterial peptidoglycan, leading to a covalently linked cell wall protein. While the LPxTG sequence is a strict necessity for sortase A recognition, a single glycine of the amino component is sufficient for the ligation step, requiring only a minimal modification of the N-terminal ligation fragment (Figure 1a).[4b] However, when sortase-mediated ligation is used for protein semisynthesis, where the native amino acid sequence of the target protein is ideally restored, the high selectivity of the LPxTG motif becomes a downside. We reasoned that a broader selectivity profile of sortase A would be advantageous, because fewer or no amino acid changes need to be introduced into the protein of interest. Here we report the evolution of a sortase A mutant with relaxed substrate selectivity by phage-display. Instead of the native leucine, this mutant tolerates several other residues in the first position of the sorting motif and recognizes a native sequence of histone H3, which can be exploited for traceless semisynthesis of this protein.

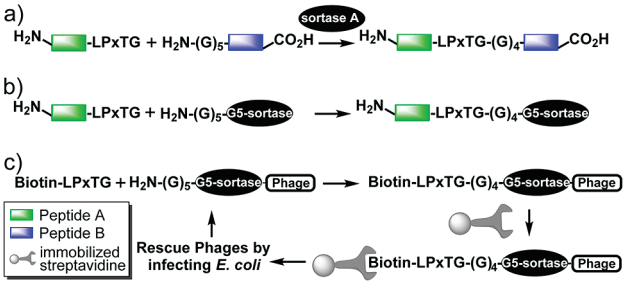

Figure 1. Directed evolution of sortase A. (a) Sortase-mediated ligation of peptides and proteins. (b) The G5-sortase catalyzes self-ligation with peptides containing the LPxTG sorting motif. (c) Product capture strategy for selecting sortase mutants possessing an altered substrate selectivity.

In the first step of this process, we developed a product capture strategy to establish a covalent bond between the enzyme and its substrate. Such a strategy prevents the accumulation of sortase mutants that hydrolyze the sorting motif, and it was successfully applied to the directed evolution of subtiligase, an engineered protease that catalyzes ligation reactions under conditions of kinetically controlled reverse proteolysis.[5] To this end, we elongated the N-terminus of *Staphylococcus aureus* sortase A by a penta-Gly-containing flexible tail (G5-sortase) (Figure 1b). G5-sortase efficiently captured a fluorescein-conjugated substrate peptide (Fl-Ahx-LPKTGGRR-NH_2, where Fl = 5(6)-carboxyfluorescein and Ahx = 6-aminohexanoic acid) by intramolecular ligation (Supporting Information, Figure S1). To ensure that the observed reaction is indeed an intramolecular product capture and not an intermolecular reaction, we additionally performed experiments under enforced intermolecular ligation conditions. Therefore, we created a G5-sortase C184A mutant which is catalytically inactive but still serves as a ligation partner for the Fl-Ahx-LPKTGGRR-NH_2 peptide when wild-type (wt) sortase without the N-terminal extension is added. When competing free penta-Gly peptide was added to either G5-sortase or G5-sortase-C184A/

Received: June 17, 2011
Published: October 06, 2011

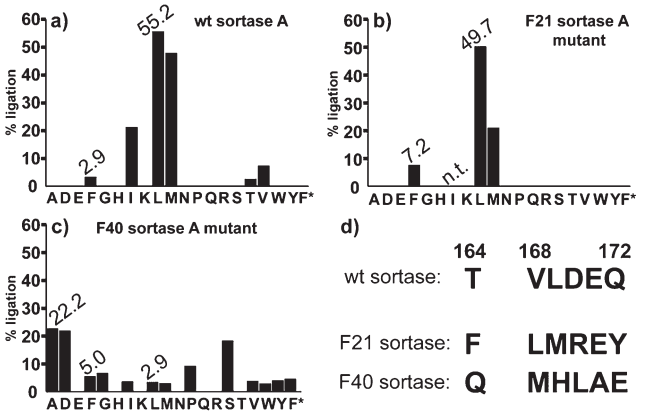

Figure 2. Characterizing the substrate preferences of the first residue in sorting motif (*X*-PKTG) of evolved sortase. Ligation reactions with a sortase motif library were analyzed for product formation by LC-MS. Product formation yields of identified ligation products were determined using individual substrate peptides. At ~55% the ligation reactions have reached equilibrium. Reaction profiles of (a) wt-sortase, (b) F21-sortase, and (c) F40-sortase are shown. (d) Amino acid sequence of the randomized $\beta6-\beta7$ loop region of selected sortase mutants. n.t. = not tested; F* = Dns-FAKTGGRR-NH_2.

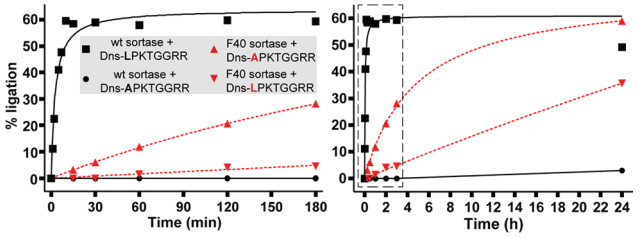

Figure 3. Comparison of the ligation of preferred substrates by wt- and F40-sortase (equimolar concentrations). The left graph highlights the difference in reaction kinetics (time scale in minutes), indicating the reduced ligation efficiency of the F40 mutant. The right graph includes a time point at 24 h showing that equilibrium for ligation of the preferred substrates is reached. After 24 h, slow hydrolysis of the LPxTG motif by wt-sortase A is observed under these conditions, resulting in decreased ligation yields for the last data point.

wt-sortase, we observed efficient self-ligation of the G5-sortase but no ligation *in trans* with G5-sortase-C184A/wt-sortase (Figure S1). These experiments showed that the reaction *in cis* is clearly favored over the ligation *in trans*. In the following we establish monovalent display of the G5-sortase as pIII fusion protein on the envelop of the M13 phage (Figure 1c). Western blot analysis of the ligation reaction with a biotinylated substrate peptide (biotin-GLPKTGGRR-NH_2) confirmed functional presentation of the G5-sortase (Figure S2).

Having established a phage-display system for selecting evolved sortases, we set out to construct a library of sortase mutants (Figures S3 and S4). Initially, we intended to screen for sortases which recognize an FPxTG sorting motif instead of the native LPxTG sequence. The reasons for choosing this variation were two-fold: on the one hand bioinformatics approaches indicated that FPxTG sorting sequences exist in nature (although the corresponding sortases remain unknown), and on the other hand we observed marginal ligation of this motif catalyzed by wt-sortase (Figure 2a).[4a] Library design was based on the crystal structure of a C184A mutant of sortase A in complex with a natural substrate.[6] Six amino acids in the $\beta6-\beta7$ loop, representing solvent-exposed positions in spatial proximity to the leucine of the sorting motif, were randomized by using NNK incorporation into synthetic oligonucleotides (Figure S4b).[7] A library of a complexity of ~10^8 was subsequently screened for the ligation of the biotin-GFPKTGGRR-NH_2 peptide to the phage-displayed G5-sortase (Figure 1c). After three rounds of panning, individual clones were sequenced and analyzed by Western blot for ligation of the biotinylated peptide (Figure S5).

Four G5-sortases showing enhanced ligation activity toward the FPxTG motif were subcloned and expressed without the N-terminal G5 extension and C-terminal pIII fusion. The substrate selectivity of these mutants was studied with an HPLC/MS-based ligation assay and a library of substrate peptides, which covered 23 variations of the sorting motif, mostly at position 1, and an alanine scan of the remaining three invariant residues (Figure S6 and Table S2). All tested sortase mutants tolerated Phe as well as other amino acids as the first residue of the sorting motif. Two mutants (F21 and F40) were further characterized in quantitative terms for the individual peptides that were identified as substrates in the library screen. The F21 mutant displayed a 2.5-fold higher ligation efficiency with respect to the FPxTG-motif than wt-sortase, but the native LPxTG remained the favored substrate of this mutant (Figure 2b). In contrast, the F40 mutant preferred ligation of the FPxTG motif over the native LPxTG sequence, but the overall ligation efficiency was low for both motifs. In general, the F40-sortase possessed a remarkable broad selectivity and accepted aromatic amino acids as well as residues with small side chains, including Ala, Asp, Ser, Pro, and Gly, at position 1 of the sorting motif (Figure 2c and Figure S7). Based on the measured reaction profile, Ala is among the most preferred amino acids in the first position of the F40-sorting motif. To assess the ligation yield that can be obtained with the F40-sortase, we monitored the ligation reaction with a Dns-APKTGGRR-NH_2 peptide (where Dns = dansyl) over a time course of 24 h (Figure 3). A final equilibrium level of ~55% was reached after 24 h. For the wt-sortase, in contrast, we detected no significant ligation of the APxTG peptide under comparable conditions, while the equilibrium level for the ligation of a LPxTG peptide catalyzed by wt-sortase is reached within 20 min (Figure 3). These observations showed that the F40-sortase has lost activity compared to the wt enzyme but is able to produce similar ligation yields over an extended period of time.

The mutations selected in the F40-sortases varied largely from the wt sequence of the $\beta6-\beta7$ loop (Figure 2d). A simple explanation for the enhanced promiscuity is not evident from the current sortase A structures, but we noticed an E171A mutation in the F40-sortase. The E171 residue is involved in calcium binding, and we therefore analyzed whether a loss of Ca^{2+} coordination caused the promiscuity of the F40 mutant. Such loss of Ca^{2+} binding conceivably increases the flexibility of the long $\beta6-\beta7$ loop and thereby allows for the recognition of a wider variety of substrates.[8] As expected, in the absence of calcium we did not observe any ligation activity of the wt-sortase with a Dns-APKTGGRR-NH_2 peptide, which is preferentially ligated by the F40-sortase mutant (Figure S8a,b). More importantly, we observed that the F40 mutant maintained Ca^{2+} dependency when ligating the Dns-APKTGGRR-NH_2 peptide, probably because the neighboring E172 mutation restores the Ca^{2+} binding site (Figure S8c). This observation indicates that the selected

dx.doi.org/10.1021/ja205630g |*J. Am. Chem. Soc.* 2011, 133, 17536–17539

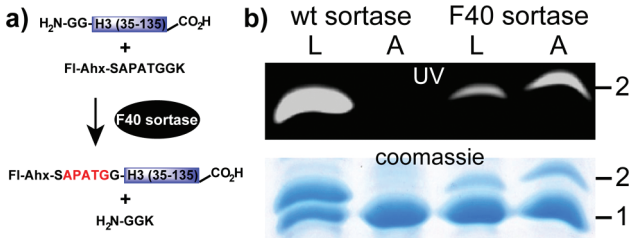

Figure 4. Semisynthesis of histone H3. (a) Scheme of histone H3 semisynthesis catalyzed by the F40-sortase. The native sequence of H3 is restored at the ligation site (marked in red). (b) Fl-conjugated peptides containing either the LPxTG or the native H3 APATG sequence were ligated to recombinant histone H3 (33−135). The unligated H3 (33−135), 1, and the ligation products, 2, were separated on SDS-PAGE and analyzed by Coomassie Brilliant Blue staining (bottom panel) and UV irradiation (top panel). L = LPxTG peptide (Fl-Ahx-LPKTGGRR-NH_2); A = APxTG peptide (Fl-Ahx-SAPATGGK-NH_2).

amino acid sequence of the $\beta6-\beta7$ loop and not the loss of Ca^{2+} binding causes the promiscuity of the F40 mutant.

We next analyzed the human proteome for traceless ligation sites of the F40-sortase mutant. The most interesting sequence was the SAPATGG site in histone H3 covering residues 28−34. Histones package DNA into chromatin and regulate gene activity through a multitude of posttranslational modifications on their N-terminal tail regions.[9] Several ligation approaches have already been tried in order to introduce site-specific modifications into the tail region of histones for biological investigations.[10] The F40-sortase mutant appeared ideally suited for such an undertaking because the APATG motif is located at the interface between the tail and the globular domain of H3, and the Ala residue in position 1 of the sorting motif is preferred by the F40 mutant. We subcloned and expressed a truncated version of *X. laevis* H3 covering residues 33−135. Two Fl-conjugated peptides containing either the sorting motif (Fl-Ahx-LPKTGGRR-NH_2) or the native 28−34 sequence of histone H3 (Fl-Ahx-SAPATGGK-NH_2) were studied in ligation assays (Figure 4a). Wt-sortase efficiently ligated the LPxTG but not the APxTG peptide to H3 (33−135). In contrast, F40-sortase catalyzed the ligation of the LPxTG only with marginal ligation efficiency, but was able to ligate the APxTG to H3 (33−135), and thereby restored the native amino acid sequence of H3 at the ligation site (Figure 4b). Encouraged by these findings, we performed a semisynthesis reaction to generate full-length histone H3. A peptide covering residues 1−33 of histone H3 was synthesized and subsequently ligated to H3 (33−135) by the F40-sortase. Western blot analysis of the ligation reaction confirmed that F40-sortase catalyzed the semisynthesis of full-length histone H3 (Figure S9).

In summary, we have reported the evolution of sortase A mutants with modified substrate selectivities. Not surprisingly, the F40-sortase mutant lost activity compared to the wt enzyme, since our screen did not select for enhanced ligation efficiency. However, the good solubility and stability of the engineered sortase allowed for efficient ligation within 24 h of reaction time. Therefore, the F40-sortase mutant can be used as a versatile tool for in vitro applications for which low enzymatic efficacy is not limiting. To the best of our knowledge, this investigation represents the first example of an engineered sortase A that maintained ligation activity when the substrate recognition site was manipulated. Earlier attempts to alter the substrate recognition of sortase A by exchanging the $\beta6-\beta7$ loop with that of sortase B, which recognizes a different sorting motif, resulted in mutants that possessed poor activity and hydrolyzed their substrates.[11] Furthermore, the engineered F40-sortase was able to catalyze traceless ligations of histone H3. This gain of function represents a very useful feature for chromatin biochemistry, enabling simple access to semisynthetic histones with defined modification patterns. In a very recent publication, it was further shown that sortase A can be evolved for improved catalytic activity.[12] Together these investigations demonstrate amenability of sortase A to engineering approaches, inspiring future manipulation experiments focused on the substrate selectivity and catalytic properties of this class of enzymes.

■ ASSOCIATED CONTENT

Supporting Information. Experimental procedures for solid-phase peptide synthesis, construction and expression of sortase constructs, phage-display, library design, ligation assays, and additional experimental data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
beyerman@fmp-berlin.de; freund@fmp-berlin.de; schwarzer@fmp-berlin.de

■ ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft to D.S. (SCHW 1163/3-1) and C.F. (SFB958, project A07). We thank Rebecca Klingberg, Till Teschke, and Bernhard Schmikale for support with solid-phase peptide synthesis.

■ REFERENCES

(1) (a) Hackenberger, C. P. R.; Schwarzer, D. *Angew. Chem., Int. Ed.* **2008**, *47*, 10030–10074. (b) Kent, S. B. H. *Chem. Soc. Rev.* **2009**, *38*, 338–351. (c) Muralidharan, V.; Muir, T. W. *Nat. Methods* **2006**, *3*, 429–438. (d) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 6974–6998.

(2) (a) Clancy, K. W.; Melvin, J. A.; McCafferty, D. G. *Biopolymers* **2010**, *94*, 385–396. (b) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707–708. (c) Popp, M. W.; Ploegh, H. L. *Angew. Chem., Int. Ed.* **2011**, *50*, 5024–5032. (d) Pritz, S.; Kraetke, O.; Klose, A.; Klose, J.; Rothemund, S.; Fechner, K.; Bienert, M.; Beyermann, M. *Angew. Chem., Int. Ed.* **2008**, *47*, 3642–3645. (e) Tsukiji, S.; Nagamune, T. *Chembiochem* **2009**, *10*, 787–798.

(3) (a) Ton-That, H.; Marraffini, L. A.; Schneewind, O. *Biochim. Biophys. Acta* **2004**, *1694*, 269–278. (b) Ton-That, H.; Mazmanian, S. K.; Alksne, L.; Schneewind, O. *J. Biol. Chem.* **2002**, *277*, 7447–7452.

(4) (a) Kruger, R. G.; Otvos, B.; Frankel, B. A.; Bentley, M.; Dostal, P.; McCafferty, D. G. *Biochemistry* **2004**, *43*, 1541–1551. (b) Mao, H. Y.; Hart, S. A.; Schink, A.; Pollok, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671.

(5) Atwell, S.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9497–9502.

(6) Zong, Y.; Bice, T. W.; Ton-That, H.; Schneewind, O.; Narayana, S. V. *J. Biol. Chem.* **2004**, *279*, 31383–31389.

(7) Bentley, M. L.; Lamb, E. C.; McCafferty, D. G. *J. Biol. Chem.* **2008**, *283*, 14762–14771.

(8) (a) Ilangovan, U.; Iwahara, J.; Ton-That, H.; Schneewind, O.; Clubb, R. T. *J. Biomol. NMR* **2001**, *19*, 379–380. (b) Ilangovan, U.;

Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6056–6061.

(9) (a) Jenuwein, T.; Allis, C. D. *Science* **2001**, *293*, 1074–1080. (b) Kouzarides, T. *Cell* **2007**, *128*, 693–705.

(10) (a) Chiang, K. P.; Jensen, M. S.; McGinty, R. K.; Muir, T. W. *Chembiochem* **2009**, *10*, 2182–2187. (b) Chatterjee, C.; Muir, T. W. *J. Biol. Chem.* **2010**, *285*, 11045–11050. (c) Schwarzer, D. *J. Pept. Sci.* **2010**, *16*, 530–537. (d) Shogren-Knaak, M.; Ishii, H.; Sun, J. M.; Pazin, M. J.; Davie, J. R.; Peterson, C. L *Science* **2006**, *311*, 844–847. (e) He, S.; Bauman, D.; Davis, J. S.; Loyola, A.; Nishioka, K.; Gronlund, J. L.; Reinberg, D.; Meng, F.; Kelleher, N.; McCafferty, D. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12033–12038.

(11) Bentley, M. L.; Gaweska, H.; Kielec, J. M.; McCafferty, D. G. *J. Biol. Chem.* **2007**, *282*, 657–681.

(12) Chen, I.; Dorr, B. M.; Liu, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11399–11404.