Published in final edited form as:

Nat Chem Biol. 2021 March ; 17(3): 317–325. doi:10.1038/s41589-020-00706-1.

Laboratory Evolution of a Sortase Enzyme that Modifies Amyloid β-protein

Christopher J. Podracky[1,2], **Chihui An**[2], **Alexandra DeSousa**[3], **Brent M. Dorr**[2], **Dominic M. Walsh**[3], **David R. Liu**[1,2,4]

[1]Merkin Institute of Transformative Technologies in Healthcare, Broad Institute of Harvard and MIT, Cambridge, MA, 02142

[2]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, 021383

[3]Laboratory for Neurodegenerative Research, Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, 02115

[4]Howard Hughes Medical Institute, Harvard University, Cambridge, MA, 02138

Abstract

Epitope-specific enzymes are powerful tools for site-specific protein modification, but generally require genetic manipulation of the target protein. Here, we describe the laboratory evolution of the bacterial transpeptidase sortase A to recognize the LMVGG sequence in endogenous Aβ protein. Using a yeast display selection for covalent bond formation, we evolved a sortase variant that prefers LMVGG substrates from a starting enzyme that prefers LPESG substrates, resulting in a >1,400-fold change in substrate preference. We used this evolved sortase to label endogenous Aβ in human cerebrospinal fluid, enabling detection of Aβ with sensitivities rivaling those of commercial assays. The evolved sortase can conjugate a hydrophilic peptide to Aβ42, greatly impeding the ability of the resulting protein to aggregate into higher-order structures. These results demonstrate laboratory evolution of epitope-specific enzymes towards endogenous targets as a strategy for site-specific protein modification without target gene manipulation, and enable potential future applications of sortase-mediated labeling of Aβ peptides.

Introduction

The ability to covalently modify proteins enables researchers to effectively interrogate and perturb their biological functions. Most purely chemical methods for protein labeling modify many proteins in a biological mixture and yield heterogeneous products that are difficult to characterize[1]. While technologies such as unnatural amino acid incorporation[2,3], inteins[4], small molecule-reactive peptides[5] and epitope-specific enzymes[6] enable chemo- and site-

*Correspondence should be addressed to David R. Liu: drliu@fas.harvard.edu.

Author Contributions. C.P., C.A., and D.R.L. designed the research, performed the experiments, and analyzed data. A.D. helped perform and analyze the CSF ELISA and aggregation experiments. B.D. provided the starting sortase library. D.M.W. designed and supervised the CSF ELISA and aggregation experiments. C.P. and D.R.L. wrote the manuscript. D.R.L supervised the research.

Competing Financial Interests. C.J.P., C.A., B.M.D., and D.R.L. have filed patent applications on sortase evolution.

selective modification in biological systems, they typically require genetic manipulation of the protein of interest to introduce an amber stop codon or peptide tag, potentially altering its biological properties and limiting applicability to settings in which target gene manipulation is possible. The ability to manipulate endogenous proteins in a site-specific manner would enable target labeling even in complex biological mixtures, and would be especially useful when genetic manipulation is impractical. To explore this possibility, we sought to evolve a versatile epitope-specific enzyme to recognize and covalently modify a peptide sequence natively present in a pathogenic protein.

Sortase transpeptidases are a superfamily of enzymes widely distributed throughout Gram-positive bacteria[7]. *Staphylococcus aureus* sortase A (SrtA) is responsible for attaching proteins that contain a C-terminal LPXTG sorting sequence to the cell wall[8]. The enzyme cleaves between the threonine and glycine of the sorting sequence, forming an acyl-enzyme intermediate that subsequently acylates the primary amine of the pentaglycine of the peptidoglycan[9]. SrtA shows a strong preference for its LPXTG sorting sequence[10], but studies have revealed that it will accept a variety of glycine-based (and some non-glycine) nucleophiles[11]. These properties make SrtA an attractive tool for site-specific protein modification. Indeed, SrtA has been successfully used for both C-terminal and N-terminal protein labeling, as well as protein circularization and the semi-synthesis of multi-domain proteins.[12–17]

Engineering of sortases for improved activity on both their cognate and novel substrates has been an area of active research for almost a decade[18,19]. Our group previously used yeast display and fluorescence-activated cell sorting (FACS) to improve the kinetics of SrtA on LPETG[20], and to evolve sortase variants that accept single amino acid substitutions at the second or fourth position of the recognition sequence[21]. In this study, we sought to reprogram the specificity of SrtA to covalently modify the Alzheimer's disease-associated amyloid β-protein (Aβ). The formation of Aβ plaques in the central nervous system is the hallmark of Alzheimer's disease (AD)[22]. Despite the clinical importance of Aβ, its physiological functions and its role in AD pathogenesis are not clearly understood[23–25]. The ability to modify Aβ site-specifically might help illuminate its biological role, impede Aβ plaque formation, or facilitate our understanding of AD pathogenesis. Since Aβ monomers are predominantly extracellular[26], unstructured[27,28], and contain a five-amino-acid sequence (LMVGG at residues 34–38) that shares features with sortase's native recognition sequence, sortase-mediated conjugation is an attractive strategy to achieve site-specific modification of Aβ.

Over 16 rounds of evolution we generated a sortase variant, SrtAβ, that mediates the covalent modification of Aβ peptides. We used SrtAβ to biotinylate and detect endogenous Aβ in clinical cerebrospinal fluid samples (CSF) at concentrations of 2–19 ng/mL. We also demonstrated that SrtAβ-mediated conjugation of a hydrophilic pentapeptide to Aβ42 greatly slows the initiation of detectable aggregation. This work establishes the evolution of sortase enzymes to site-specifically modify naturally occurring proteins without requiring modification of endogenous genes.

Results

Initial evolution of SrtA to recognize Aβ

We sought to evolve SrtA variants that modify Aβ using yeast display[20,29–31] and fluorescence-activated cell sorting (FACS) (Figure 1). Briefly, yeast display a library of sortase variants conjugated to triglycine peptides with N-termini that are free for sortase-catalyzed reactions. The library is then incubated with an N-terminally biotinylated target substrate and non-biotinylated off-target substrates. Sortase variants that catalyze transpeptidation between triglycine and the target substrate biotinylate the surfaces of the yeast cells that encode them. Activity on off-target substrates by promiscuous sortase variants leads to reduced biotinylation of the cells that encode them. After removal of cell surface-displayed sortases with TEV protease (Supplementary Figure 1), cells are stained with fluorophore-linked streptavidin and the biotinylated cells encoding active and selective sortase variants are isolated by FACS (Extended Data Figure 1).

We started our evolution from a library of sortase variants previously evolved to recognize LPESG substrates (library 4S.6). Given that our target sequence, LMVGG, deviates from the wild-type sorting sequence, LPXTG, at the second and fourth positions, we reasoned that mutants already possessing altered substrate recognition at the fourth position were a more promising starting point than wild-type SrtA. We diversified this starting pool by error-prone PCR to create the round 1 library of 4.8×10^7 variants. To identify variants that preferred glycine over serine at the fourth position, we used biotinylated LPVGG (Btn-LPVGG) as an initial positive selection substrate. The stringency of the screen was gradually increased by decreasing the Btn-LPVGG concentration and increasing the off-target non-biotinylated LPESG substrate concentration (Supplementary Table 1). We isolated individual clones after five cycles of enrichment. Prominent mutations from round 1 included R94P, S118I G134R, and V189F (Table 1).

We re-diversified the pool by error-prone PCR to create the round 2 library. This library showed sufficient activity on Btn-LMVGG to permit FACS using this substrate. As in round 1, stringency was increased by reducing the amount of positive selection substrate while increasing the amounts of negative selection substrates, in this case LPVGG and LMVTG (Supplementary Table 1). After round 2, we observed that the S118I, G134R, and V189F mutations from round 1 had persisted, but that the identity of residue 94 was diverse (Tyr, Leu, Arg, Pro, His, or Gln) among sequenced clones. In addition, ~98% of sequenced clones had mutations at residue 124 (Asp to Gly, Leu, or Tyr).

Rounds 3 through 7 consisted of iterative cycles of diversification by error-prone PCR and FACS screening for activity on Btn-LMVGG with progressively higher stringencies (Supplementary Table 1). At the end of round 3, we observed a clone that represented 3.5% of the population and contained new mutations K138I, V182A, T196S, and R197S, in addition to previously observed mutations R94Y, S118I, D124L, G134R, and V189F. By the end of round 4 this clone represented 74% of the population (Table 1), suggesting a substantial fitness benefit from some combination of K138I, V182A, T196S, and R197S.

The most common sequence emerging from round 5 (36% of the population) was the round 4 consensus sequence plus an I123L mutation. I123L was the most common new mutation emerging in round 5, present in 67% of sequenced clones. Notably, the V182A, T196S, and R197S mutations that first appeared at the end of round 3 reached 100% prevalence in the population. Following two additional rounds of diversification and sorting, the consensus sequence of the round 7 pool (29% of the population) contained R94Y, S118I, I123L, D124L, G134R, K138I, K173E, V182A, V189F, T196S, and R197S (Table 1). Analysis of previous sequencing data showed that this clone first appeared at the end of round 5, where it made up 9% of the population.

Of these 11 mutations, we were particularly interested in V182A, T196S, and R197S because of their early prevalence. Additionally, mutations at residues 182 and 196 were previously observed in SrtA variants with improved activity or single-position altered substrate recognition[20,21], while residue 197 is a crucial part of the active site in wild-type SrtA[32–35]. We generated the round 8 library using site-saturation mutagenesis at these three positions followed by error-prone PCR. Mutations V182A, T196S, and R197S remained fixed in sequences emerging from round 8, confirming the fitness advantage afforded by these three mutations. Additional well-represented mutations that appeared in round 8 include K62R (present in 84% of sequenced clones), I76L (60%), the reversion mutation Y94R (62%), N107D (20%), N127Y (15%), I138L (15%), K145T (15%), M155I (20%), R159C (15%), K173E (57%), K177R (24%), and F189I (35%).

SrtA evolution in human plasma

While the above screens for sortase activity on LMVGG were conducted in TBS buffer, our goal of modifying Aβ in endogenous contexts requires that the evolved enzyme be active in biological fluids. Our previous work revealed that sortase enzymes evolved for LPESG recognition—including clone 4S.6, the starting point of this study—are capable of modifying fetuin A in human plasma, presumably through its native LPPAG sequence[21]. Indeed, a 4-fold molar excess of a round 8 clone also supported labeling of purified fetuin A in DPBS (Supplementary Figure 2a), and overnight incubation of human plasma with 50 μM evolved sortase and 1 mM GGGK(Btn) also led to fetuin A labeling (Supplementary Figure 2b). To evolve decreased recognition of fetuin A, we conducted additional rounds of evolution with negative selection against the LPPAG sequence of fetuin. This negative selection was achieved by including LPPAG peptide in our sortase reaction mixtures (round 9) and by conducting the sortase reactions directly in human plasma (rounds 10–16) Over these eight rounds of evolution, we generated a sortase variant with greatly reduced activity on fetuin A relative to the starting sortase 4S.6 (Supplementary Figure 2c and d).

Between the end of round 8 and the beginning of round 9, we generated and analyzed the activity of a series of single-reversion mutants from the round 7 consensus sequence. These data (Supplementary Figure 3) revealed the importance of mutations at residues 94, 123, and 124. As such, we conducted site-saturation mutagenesis at these three residues and adjacent residue 122, followed by error-prone PCR to generate the round 9 library. Increased off-target LPPAG concentration and decreased reaction times were used to increase selection stringency over the course of round 9 screening (Supplementary Table 1). Sequencing the

pool at the end of round 9 revealed enrichment of many mutations that we first observed in round 8. This included K62R (up to 100% from 84% of sequenced clones), I76L (up to 97% from 60%), the reversion mutation Y94R (up to 91% from 62%), N107D (up to 61% from 20%), I138L (up to 42% from 15%), K145T (up to 55% from 15%), R159C or H (up to 21% and 53% from 15% and 4%, respectively), K173E (up to 85% from 57%), K177R (up to 72% from 24%), and F189I (up to 72% from 35%) (Table 1).

To maintain selection against fetuin A recognition while introducing selection against other motifs that exist in human plasma, we conducted the sortase reactions for our screens directly in human plasma from round 10 onward. 100-fold higher concentrations of Btn-LMVGG were initially needed to observe sortase conjugation in human plasma than were needed to observe conjugation in TBS, suggesting that specific labeling of the desired target is more difficult in plasma. (Supplementary Table 1). Analysis of the round 10 sequencing results showed further enrichment of N107D (present in 95% of sequenced clones), N127Y (56%), M155I (35%), and K145T (98%). The Y94R reversion, I138L, R159C, and K173E all reached 100% abundance by the end of round 10. New mutations included N127H (27%) and Q172H (65%). Round 11 resulted in further enrichment of N127H (present in 75% of sequenced clones), M155I (70%), Q172H (100%), and the appearance of the G139D (28%). N172H and M155I further enriched in round 12 (both to 87% abundance), where an E105D mutation appeared in 32% of clones. By the end of round 13, E105D was found in 88% of sequenced clones.

Our inability to use more stringent conditions in round 14 than in round 13 (Supplementary Table 1), coupled with the low convergence of the resulting pools, prompted us to use DNA shuffling in an attempt to escape a potential fitness plateau. We shuffled the round 14 pool with the eSrtA pentamutant in a 1:1 ratio and subjected the products to error-prone PCR to create the round 15 library. In round 15 we used Aβ40 conjugated to biotin at its N-terminus through an aminohexanoic acid linker (Btn-LC-Aβ40) as the target substrate instead of Btn-LMVGG to ensure activity on the full target peptide and not only the recognition motif. In round 16 we used a 1:1 mixture of Btn-LC-Aβ40 and Btn-LC-Aβ42 to select for activity on two different Aβ alloforms. The most notable mutation to emerge in round 16 was S102C, which was present in all four of the most active individual clones. Other mutations of note include M141I, K152R, and K206R. Given the level of activity already observed and the lack of additional strongly enriching mutations after round 13, we ended the evolution campaign and characterized the evolved sortase enzymes.

Flow cytometry analysis of the pools at the end of each screening round revealed an upward trend in activity on the LMVGG substrate (from round 1 through round 9, Extended Data Figure 2). An initial downward trend in pool activity was observed upon switching to human plasma as reaction buffer. This downward trend was reversed in round 12, but activity dropped and plateaued again in rounds 13 and 14. The noticeable increase in activity from round 14 to round 15 both in TBS and in plasma (43% increase in TBS, 54% in plasma) suggests that DNA shuffling was a successful strategy to escape an apparent fitness plateau, and the overall trend in strongly increased activity between round 1 and round 16 confirmed a successful evolutionary campaign for a SrtA variant with activity on Aβ.

Characterization of key mutants and mutational analysis

At the end of round 8, individual variants were isolated by sorting single cells into a 96-well plate. The activity of 32 evolved sortase clones towards LMVGG and LPVGG was assessed by a flow cytometry assay (Extended Data Figure 3). Clone 8.5-H3 demonstrated the best combination of activity on LMVGG and selectivity over LPVGG. When expressed and purified, this variant was active on LMVGG in an established HPLC assay for SrtA activity[36], converting 10% of 10 μM substrate to product in two hours, an improvement from previous rounds (Extended Data Figure 4). Kinetic parameters for 8.5-H3 were determined to be k_{cat} = 0.012 s^{-1} (95% CI = 0.009 to 0.017 s^{-1}) to with K_M = 52 μM (95% CI = 29 to 103 μM) using an established fluorescence assay[36,37]. Western blot analysis revealed that variant 8.5-H3 was able to conjugate a variety of Aβ isoforms with GGGK(Btn), demonstrating that sortases evolved to process LMVGG also show activity on Aβ (Extended Data Figure 5).

Individual variants from round 16 were sorted and re-assayed for LMVGG activity at the end of the round. The top variant from round 16 (SrtAβ) was assayed by flow cytometry on a panel of substrates, which revealed a greatly altered substrate profile from the starting enzyme 4S.6 (Figure 2a). The results were consistent with positive selection for activity on LPVGG in round 1 and LMVGG in subsequent rounds with negative selection against LPESG in round 1 and against LPPAG in round 9–16. Compared to the starting enzyme 4S.6, SrtAβ has 53-fold reduced activity on LPESG, 11-fold reduced activity on LPPAG, and 28-fold increased activity on LMVGG (Figure 2a). SrtAβ has a 30-fold preference for LMVGG over LPESG, whereas 4S.6 has a 49-fold preference for LPESG over LMVGG. Overall, SrtAβ evolved a 1,470-fold change in preference to favor LMVGG over LPESG.

SrtAβ contains 25 mutations relative to the starting sortase enzyme 4S.6. To determine the relative importance of individual mutations to activity on LMVGG, we reverted each mutation back to its corresponding residue in the starting enzyme. These 25 single-mutant variants were assayed alongside SrtAβ and 4S.6 by flow cytometry (Figure 2b). Eleven of the reversions reduced enzyme activity by less than 25%, eight reduced activity between 25–50%, and six reduced enzyme activity by at least two-fold. Notably, reversion mutations at residues 118 and 197 reduced activity on the LMVGG substrate greater than 90%, near the low level of activity demonstrated by the starting enzyme 4S.6. Two of these six mutations are at residues that were identified as modulators of sortase substrate specificity in our previous evolution campaigns (residues 118 and 182)[21], but the remaining four were at novel residues. Notably, three of these four novel residues are outside of the substrate binding pocket, and the fourth, R197, is highly conserved across the sortase superfamily. It has been suggested that R197 in wild-type *S. aureus* SrtA stabilizes the binding of the LPXTG sorting signal or the oxyanion intermediates generated during catalysis.[35] That a non-conservative mutation at this residue is not only tolerated, but required, is surprising. These results highlight the challenge of *a priori* prediction of mutations that alter SrtA specificity, and the importance of including random mutagenesis as a diversification strategy (Figure 2c).[35] A minimal mutant containing these six mutations in the 4S.6 background showed a 4-fold improvement in LMVGG activity relative to 4S.6, but 23-fold lower activity than SrtAβ (Figure 2d). This result confirms that other mutations, though less important individually,

collectively contribute to substantially improved target activity. Four of these other mutations, in addition to R177K, are at residues located near the calcium binding site in the wild-type enzyme. Assaying SrtAβ activity at various calcium concentrations revealed compatibility with a broad range of concentrations (0.1 to 10 mM) that include physiological calcium concentrations, but confirmed that calcium is still required for activity (Supplementary Figure 4).

To confirm that the shift in substrate specificity observed in the flow cytometry assay translated to purified enzymes, we determined the kinetic parameters of 4S.6 and SrtAβ on LPESG and LMVGG using a HPLC assay (Extended Data Figure 6). Sortase 4S.6 showed k_{cat} = 0.36 s^{-1} (95% CI = 0.22 to 0.96 s^{-1}) and K_M = 610 μM (95% CI = 90 to 5550 μM) on LPESG, whereas SrtAβ activity on LPESG was too low to establish accurate kinetic parameters. SrtAβ had k_{cat} = 0.018 s^{-1} (95% CI = 0.015 to 0.023 s^{-1}) and K_M = 128 μM (95% CI = 87 to 198 μM) on LMVGG, whereas 4S.6 activity on LMVGG was not detectable. These findings confirm that the evolution resulted in a large change in substrate preference, consistent with the >1,400-fold change observed in flow cytometry assays (Figure 2a).

To obtain a more quantitative understanding of our evolved enzyme's activity on Aβ40 in plasma, we developed an ELISA to measure biotinylated Aβ. Streptavidin was used to capture biotinylated peptide and detection accomplished using 4G8, a monoclonal antibody that recognizes Aβ residues 17–24. Aβ40 labeled with GGGK(Btn) was used as the calibrant. Employing this assay, we confirmed SrtAβ activity on Aβ40 spiked into human plasma, with 1.5 μM SrtAβ generating 2.3 μM of biotinylated product from 5 μM Aβ40 in two hours (Extended Data Figure 7a). As we expected, increasing the amount of GGGK(Btn) nucleophile greatly improved reaction yields (Extended Data Figure 7b).

Concentrations of Aβ peptides are important biomarkers of Alzheimer's disease. This is especially true of Aβ42 in CSF, where a decrease to roughly 50% of baseline Aβ42 levels is typically observed in AD patients[38]. To enable labeling and detection of physiologically relevant amounts of Aβ, we changed the format of the ELISA to capture the product with monoclonal antibody m266 (the epitope of which spans Aβ residues 13–26) and detect with streptavidin-HRP. After optimizing the concentrations of various assay components, we were able to consistently detect and quantify Aβ-Btn conjugates at concentrations comparable to commercial Aβ ELISA kits (Supplementary Table 2). The lower limit of quantitation (LLoQ) is defined as the lowest standard with a signal higher than the average signal of the blank samples plus nine standard deviations, and allows a percent recovery of 80–120%. In six runs over six days, LLoQ for our assay was 39–78 pg/mL, or roughly 10–20 pM. Using this SrtAβ-mediated assay, we observed labeling of Aβ40 spiked into human plasma at concentrations as low as 5 nM (Extended Data Figure 7c). Given that typical Aβ concentrations in human CSF are on a similar order of magnitude, these observations suggest the possibility of using SrtAβ to label endogenous Aβ in CSF, where the generation, clearance, and aggregation of Aβ are all intimately connected with AD etiology[24,39].

SrtAβ labels endogenous Aβ in CSF

The ability to site-specifically modify endogenous Aβ in CSF would provide researchers with new ways to interrogate or influence these dynamic processes. Thus, we sought to demonstrate labeling of endogenous Aβ in CSF. First, we measured Aβ levels in CSF samples using immunoassays specific for Aβ terminating at Val40 or Ala42[40]. Because sortase-mediated conjugation of Aβ40 and Aβ42 destroys the C-terminal epitopes used for immunodetection, we reasoned that this reaction would cause a loss of ELISA-measured signal. Indeed, after treating the samples with SrtAβ and GGG, we observed losses in signal ranging from 47% to 77%, confirming that the enzyme was active in CSF (Figure 3a).

While this loss of signal is consistent with transpeptidation, it might also be explained by hydrolysis or interference of the sortase enzyme with the binding of the detection antibody. Loss of signal due to aggregation is unlikely since we have previously shown that incubation of biological samples at room temperature for up to 24 hours does not alter detection of Aβ42[41]. Besides the enzyme and GGG, the only difference between treated and untreated samples was the addition of calcium, a cation known to influence in vitro aggregation of Aβ[42]. However, CSF already contains micromolar levels of calcium and it is unlikely that a modest increase in calcium would induce aggregation of Aβ present at nanomolar concentrations.

To obtain a more direct read out of transpeptidation activity, we generated AβM1–37-GGGK(Btn) semi-synthetically (Supplementary Figure 5a), and used it as our standard to detect reaction product generated by SrtAβ-catalyzed conjugation with GGGK(Btn). As before, Aβ peptides were captured using the m266 antibody and detected via streptavidin-HRP. We observed Aβ labeling efficiencies of 13% to 56% (Figure 3b). These efficiencies are lower than those observed with GGG labeling, as expected given that only transpeptidation leads to gain of signal in this assay. The lower efficiency is not likely due to SrtAβ having a preference for GGG over GGGK(Btn), since reactions of chemically synthesized Aβ with equimolar amounts of different triglycine nucleophiles yield similar amounts of transpeptidation products (Supplementary Figure 5b). The variable labeling efficiencies across samples suggests that this method requires further optimization for use in absolute quantification of Aβ. Nonetheless, these data demonstrate the ability of the evolved SrtAβ enzyme to modify endogenous Aβ in human CSF.

Sortagging Aβ42 alters aggregation kinetics

Next, we sought to conjugate Aβ to a molecule that would impede its aggregation. Previous studies showed that the hydrophobic C-terminus of Aβ42 is well-resolved in the NMR solution structure of Aβ42 fibrils[43]. We hypothesized that the replacement of hydrophobic C-terminal residues with more hydrophilic residues would alter the aggregation propensity of the resulting peptides.

To test this possibility, we expressed and purified Aβ42 as previously reported[44]. Immediately following batch purification, we treated a portion of the recombinant Aβ42 (20 mL of ~40 μM) overnight with 20 μM SrtAβ and 200 μM GGGRR. Transpeptidation should replace the last five residues of Aβ(M1–42), GVVIA, with GGGRR, yielding a more

hydrophilic 43-mer. As expected, Aβ(M1–37-GGGRR), the identity of which was confirmed by mass spectrometry, eluted from reverse-phase HPLC before AβM1–42 (Supplementary Figure 5c). We then directly compared the aggregation propensity of the HPLC-isolated Aβ(M1–37-GGRR) to that of recombinant AβM1–42 from the same initial batch purification.

Using a continuous thioflavin T (ThT) binding assay[45], the SrtAβ-modified peptides were found to take much longer to nucleate into aggregates. The lag time to initiation of detectable aggregation for 20 μM Aβ(M1–42) occurred within 5 minutes, whereas the lag time for 20 μM Aβ(M1–37GGGRR) was 8.2 hours. The modified peptides also took ~40-fold longer to reach half maximal aggregation (0.70 or 0.56 hours for 10 or 20 μM AβM1–42 versus 28 or 14.6 hours for 10 or 20 μM Aβ(M1–37GGGRR)) (Figure 4). The impaired aggregation of SrtAβ-modified AβM1–42 was replicated with recombinant Aβ(M1–37GGGRR) (Extended Data Figure 8). In contrast to the delayed kinetics of aggregation, the maximum ThT signals of C-terminally modified fibrils were higher (46,000 vs 14,000 RFU at 20 μM, 30,000 vs 5,000 RFU at 10 μM). Thus, while the lag time for AβM1–37GGGRR was much longer than for AβM1–42, the rate of aggregation and the extent of ThT binding was greater for AβM1–37GGGRR. These results indicate that modification of the Aβ C-terminus delays nucleation, but once nuclei are formed elongation is rapid and the structure formed binds ThT in a manner distinct from Aβ42. Indeed, EM analysis of aggregation end-products revealed substantial ultrastructural differences in the fibrils formed by AβM1–37GGGRR and AβM1–42 (Extended Data Figure 9). Collectively, these results establish the modification of a disease-associated form of Aβ to a form less prone to aggregation by transpeptidation using a laboratory-evolved sortase enzyme.

Discussion

We used a yeast display selection strategy over many rounds of evolution to generate a sortase enzyme capable of site-specifically modifying Aβ peptides. We leveraged the ability to tailor reaction conditions during selection by lowering the concentration of target substrate, altering the kinetic requirements to survive selection, and introducing various decoy off-target substrates, thereby tuning selection stringency for activity and specificity. After 16 total rounds of evolution, we generated a sortase variant that prefers LMVGG 30-fold over LPESG, a large change in specificity from a starting enzyme that prefers LPESG 49-fold over LMVGG. To our knowledge, this work represents the first example of a sortase enzyme evolved for activity on a substrate with mismatches at multiple amino acids in its recognition sequence.

Previous efforts to reprogram sortase activity yielded orthogonal variants that were highly active on singly mutated LAXTG and LPXSG substrates[21]. The present results demonstrate that it is possible to evolve an epitope-specific enzyme capable of recognizing an endogenous peptide tag in a disease-associated protein (LMVGG, residues 34–38 of Aβ), although the catalytic efficiency of the evolved enzyme (143 $M^{-1}s^{-1}$) on this new substrate is lower than evolved to recognize the singly mutated targets (~10^3–10^4 $M^{-1}s^{-1}$)[21]. This difference in efficiency largely arises from a lower k_{cat}, as the K_M of SrtAβ for LMVGG (128 μM) is improved relative to the K_M of 4S.6 for LPESG (610 μM) and falls within the

range of K_M values reported for other sortase variants evolved by yeast display[20,21]. In our current scheme, each copy of a given library member is only allowed a single turnover with which to generate signal. Development of a multiple turnover variant of this selection for bond-forming enzymes, perhaps by displaying two sets of Aga2p-fusions on the cell surface[46], could facilitate further improvements in turnover number and catalytic efficiency. The identification of novel peptide ligases as evolutionary starting points could also broaden access to targets that otherwise would be inaccessible to sortase A[47–49].

We demonstrated the ability of our evolved enzyme, SrtAβ, to generate conjugates with Aβ monomers, validating the evolution of epitope-specific enzymes as a strategy for site-specific labeling of endogenous peptides. An enzyme capable of site-specific Aβ modification enables a variety of applications. We successfully generated conjugates of purified Aβ with GGGK(Btn) and GGGRR, among other peptides. The ability of sortase enzymes to use a wide array of glycine-based nucleophiles means that the applications of our evolved enzyme with purified Aβ monomers are not limited to those described herein. For example, SrtAβ could be used to generate novel Aβ-adjuvant conjugates for vaccine development efforts[50].

SrtAβ can also conjugate peptides to endogenous Aβ in human CSF, raising the possibility of biomedical applications on endogenous Aβ. Attachment of fluorophores could enable imaging studies that further our understanding of AD etiology. Tagging monomers to inhibit their aggregation, as demonstrated above, or to mark them for degradation, could modify disease state. As such, SrtAβ or other variants could help illuminate the biological role of Aβ, increase our understanding of AD pathogenesis, and potentially contribute to the development of new AD treatments.

Online Methods

Library diversification by error-prone PCR

Genes were isolated from harvested yeast libraries by PCR using the primers pCTCon2CTEV.HR2.Fwd and pCTCon2CTEV.HR2.Rev, purified by gel electrophoresis, and subsequently mutagenized by using the GeneMorph II Random Mutagenesis Kit (Agilent) for 25 cycles of PCR amplification using primers pCTCon2CTEV.HR2.Fwd and pCTCon2CTEV.HR2.Rev. Reactions were purified by spin column and combined with *Nhe*I/*Bam*HI-digested pCTCon2CTev vectors in a 5:1 insert:backbone mass ratio and electroporated into ICY200 as described below to yield yeast libraries.

pCTCon2CTEV.HR2.Fwd: CCCATACGACGTTCCAGACTATGCAGGATCTGAG AACTTGTACTTTCAAGGTGCT

pCTCon2CTEV.HR2.Rev: CTGTTGTTATCAGATCTCGAGCTATTACAAGTCCT CTTCAGAAATAAGCTTTTGTTCGGA

Library diversification by site saturation mutagenesis (rounds 8 and 9)

Genes were isolated from harvested yeast libraries by PCR using the primers pCTCon2CTEV.HR2.Fwd and pCTCon2CTEV.HR2.Rev, purified by gel electrophoresis,

and subcloned into pET29 via restriction digest with *Nhe*I/*Bam*HI. This plasmid was used as the template for site-saturation mutagenesis with PNK-treated primers:

182-NNK-Fwd: NNKACCTGCGATGATTATAACTTTGAAACCG

182-NNK-Rev: CAGGGTCAGCTGTTTATCTTTGCC

196–197-NNK-Fwd: NNKNNKAAAATTTTTGTGGCGACCGAAGTG

196–197-NNK-Rev: TTCCCACACGCCGGTTTC

in round 8 and:

94-NNK-Fwd-1: NNKGAACAGCTGGATCGTGGCGTGAGC

94-NNK-Fwd-2: NNKGAACAGCTTGATCGTGGCGTGAGC

94-NNK-Rev: GGTCGCCGGGCCCGG

122–124-NNK-Fwd-1:
NNKNNKNNKCGTCCGAACTATCAGTTTACCAACCTG

122–124-NNK-Fwd-2: NNKNNKNNKCGTCCGTACTATCAGTTTACCAACCTG

122–124-NNK-Rev: GGTATGGCCGATAATGCTAATGTTCTGATC

in round 9. Site-saturated genes were then amplified out of the pET29c backbone using primers pCTC-HR-pET29-Fwd and pCTC-HR-pET29-Rev and purified by gel electrophoresis.

pCTC-HR-pET29-Fwd:
CCCATACGACGTTCCAGACTATGCAGGATCTGAGAACT
TGTACTTTCAAGGTGCTAGCCAGGCGAGACCGCAGATTCC

pCTC-HR-pET29-Rev:
CTGTTGTTATCAGATCTCGAGCTATTACAAGTCCTCTTC
AGAAATAAGCTTTTGTTCGGA TCCTTTCACTTCGGTCGC

Library diversification by DNA shuffling (round 15)[51]

The harvested library from the end of round 14 and the evolved sortase A pentamutant (5M) were each amplified with pCTCon2CTEV.HR2.Fwd and pCTCon2CTEV.HR2.Rev and purified by gel electrophoresis. 1 μg of each PCR product was added to 5 μL of 500mM Tris-HCl pH 7.4, 100mM $MnCl_2$ and brought to 50 μL total volume. This mixture was incubated at 15°C for 5 minutes at which point 0.5U of DNaseI was added. After 90 seconds, 1 μL of 500 mM EDTA was added to the reaction and the enzyme was heat killed at 90°C for 10 minutes. The digest was run on a 3% agarose gel and 25–150 bp fragments were isolated. 200 ng of DNA fragments were added to a 100 μL primerless reassembly reaction with 5 μL 4 mM dNTPs, 4 μL 50 mM $MgSO_4$, 10 μL 600 mM Tris-SO_4 (pH 8.9)/180 mM ammonium sulfate, 1U Taq polymerase, and 1U Phusion polymerase. This reaction was cycled at 94°C for 2min, then 35 cycles of (94°C for 15sec, 65°C for 45sec, 62°C for 45sec, 59°C for 45sec, 56°C for 45sec, 53°C for 45sec, 50°C for 45sec, 47°C for 45sec, 44°C for 45sec, 41°C for 45sec, 68°C for 45sec), and then 68°C 1min. After PCR

cleanup, a portion of the primerless reassembly product was amplified with primers CJP66-Fwd and CJP66-Rev, digested with *Nhe*I/*Bam*HI and ligated into pCTCon2CTev vector.

CJP66-Fwd: GTACTTTCAAGGTGCTAGCC

CJP66-Rev: CAGAAATAAGCTTTTGTTATC

Yeast library construction

Fresh plates of ICY200 *S. cerevisiae* cells were streaked from long-term glycerol stocks and grown for 72 hours at 30 °C prior to use. A single colony was picked and grown in 10 mL YPD + 100 U/mL penicillin, 100 μg/mL streptomycin, 100 μg/mL kanamycin overnight with shaking at 30 °C. This suspension culture was freshly diluted into 125 mL YPD and electrocompetent cells were prepared as described by Chao et al.[52] All library transformations were performed by gap repair homologous recombination into pCTCon2CTev vectors linearized by *Nhe*I and *Bam*HI digestion. Following transformation, 10^5 and 10^6 dilutions were plated and used to estimate library size.

Yeast library induction

Libraries were grown in SCD-Trp-Ura dropout media + 100 U/mL penicillin, 100 μg/mL streptomycin, 100 μg/mL kanamycin at 30 °C. Library expression was induced by transfer to SGR-Trp-Ura media at 20 °C overnight.

GGGK-CoA Synthesis

Fmoc-GGGK-$CONH_2$ was dissolved in DMSO to a final concentration of 100 mM, then combined with 1.5 equivalents of LC-SMCC (Thermo-Fisher) and 2 equivalents of DIPEA (Sigma) in DMSO. The reaction was incubated for 1 hr at room temperature, then combined with 1.5 equivalents of coenzyme A trilithium hydrate (Sigma) in DMSO to a final peptide concentration of 25 mM and mixed at room temperature overnight. The Fmoc protecting group was removed with 20% vol/vol piperidine and incubation for 20 minutes. The reaction was quenched by the addition of 1 equivalent of TFA, and the product was purified on a preparative Kromasil 100–5-C18 column (21.2×250 mm, Peeke Scientific) by reverse phase HPLC (flow rate: 9.5 mL/min; gradient: 10% to 70% acetonitrile with 0.1% TFA in 0.1% aqueous TFA gradient over 30 minutes; retention time: 17.1 minutes). ESI-MS: $[M-H]^-$ *m/z* = 1300.1 (observed); calculated for $C_{45}H_{72}N_{14}O_{23}P_3S^-$ = 1301.4. The concentration of GGGK-CoA peptide was determined from the measured A_{259} using the known molar extinction coefficient of coenzyme A [53], 15,000 M^{-1} cm^{-1}.

Sfp expression and purification

E. coli BL21(DE3) harboring the pET29 expression plasmid for Sfp phosphopantetheinyl transferase were cultured at 37 °C in LB with 50 μg/mL kanamycin until OD_{600} ~0.6. IPTG was added to a final concentration of 1 mM, and protein expression was induced at 37 °C for three hours. The cells were harvested by centrifugation and lysed by resuspension in B-PER(Novagen) containing 260 nM aprotinin, 1.2 μM leupeptin, 2 units/mL DNAseI, and 1 mM PMSF. The clarified supernatant was purified on Ni-NTA agarose, and fractions that were >95% pure were consolidated and dialyzed against 10 mM Tris pH 7.5 +1 mM EDTA

+5% glycerol. Enzyme concentration was calculated from the measured A280 using the published extinction coefficient of 27,220$M^{-1}cm^{-1}$.[54]

TEV protease expression and purification[55]

E. coli BL21(DE3) harboring the pRK793 plasmid for TEV S219V expression and the pRIL plasmid (Addgene) was cultured in LB with 50 μg/mL carbenicillin and 30 μg/mL chloramphenicol until OD_{600} ~0.7. IPTG was added to a final concentration of 1 mM, and the cells were induced for three hours at 30 °C. The cells were pelleted by centrifugation and lysed by sonication. The clarified lysate was purified on Ni-NTA agarose, and fractions that were >95% TEV S219V were consolidated and dialyzed against TBS. Enzyme concentrations were calculated from A_{280} measurements using the reported extinction coefficient of 32,290 M^{-1} cm^{-1}.

Yeast library preparation and fluorescence-activated cell sorting

Induced cells were pelleted and resuspended in 10 mL TBS-B (100 mM Tris pH 7.5, 500 mM NaCl, 1% BSA). To this cell suspension was added 50 μL 1 M $MgCl_2$, 10 μL 200 mM H_2NGGGK(CoA), and 50 μL 100 μM Sfp (10 mM Tris pH 7.5, 1 mM EDTA, 10% glycerol). The Sfp ligation reaction was incubated for 45 min at room temperature. Cells were then pelleted at 2400 g x 10 min and the supernatant was removed. Desired sortase reaction buffer (TBS-BC; 100 mM Tris pH 7.5, 500 mM NaCl, 1% BSA, 5 mM $CaCl_2$, or PC; human plasma (GeneTex, GTX73265) centrifuged at 21000 g x 10 min and passed through a 0.4 micron filter, 5 mM $CaCl_2$) was then added and the cell pellet resuspended.

Separately, 100x target substrate and negative selection substrates (custom syntheses from Genscript, with the exception of Btn-LC-Aβ40 and Btn-LC-Aβ42 obtained from ERI Amyloid Laboratory, LLC) were added to Eppendorf tubes. Typically, this involved 3–4 aliquots of varying substrate concentration such that a range of selection stringencies is represented across the aliquots. Cell suspension was added to the substrates, inverted to mix, and incubated for 15 to 60 min at room temperature. Cells were pelleted and treated with 1 mL TEV solution (100 μg/mL in PBS, 0.5% BSA, 2 mM EDTA) for 30 min on ice. Cells were pelleted and labeled with antibodies (1:200 Streptavidin-PE and 1:250 anti-HA Alexafluor-488, both from Invitrogen, in PBS, 0.5% BSA, 2 mM EDTA) for at least 30 min on ice. Cells were pelleted and washed once with 1 mL PBS, 0.5% BSA, 2 mM EDTA, then suspended in the same buffer before analysis and sorting on a BD FACS Aria Cell Sorter.

A negative control lacking any biotinylated target substrate was used to draw gates for sortase activity:expression level (PE:FITC) (see Supplemental Figure 2). Aliquots that contained target substrate were then analyzed, and the number of events in the PE:FITC gate was compared to the negative control. Aliquots that showed a >10-fold increase in gated events versus the negative control were considered suitable for sorting. The top 0.5–1.0% of cells were collected from a total number of events at least 10-fold greater than the estimated library size.

Cells sorted in active gate were collected in 2 mL SDC –Trp –Ura dropout media + 100 U/mL penicillin, 100 μg/mL streptomycin, 100 μg/mL kanamycin in a 15 mL conical. Collected cells were then divided into 2 – 4 10 mL SDC –Trp –Ura cultures and grown at 30

°C for 2 days before they were induced again for a subsequent sort under more stringent conditions. Increased stringency was most commonly achieved by decreasing target substrate concentration, but occasionally by increasing off-target concentration or decreasing reaction times. Cycles of growth, induction, and enrichment were iterated until active variants could no longer be isolated using more stringent conditions than those used in the previous cycle, generally about 4–6 times. At this point, the surviving pool was extracted and re-diversified to create a library for the next round.

Yeast library harvesting

Following the final FACS screen of a round, yeast were grown to saturation (OD ~1.5) in SCD –Trp –Ura dropout media + 100 U/mL penicillin, 100 μg/mL streptomycin, 100 μg/mL kanamycin at 30 °C, then lysed using a Zymo Research Zymoprep II kit according to manufacturer's instructions

Isolation of single clones

A portion of the harvested plasmid was transformed directly into Thermo Fisher One-Shot Mach1 T1 Chemically Competent cells according to manufacturer's instructions. 36–48 colonies, each bearing a single library member, were picked for rolling circle amplification and subject to Sanger sequencing with primers:

CA205: AGGCAATGCAAGGAGTTTTTG

CA232: CAGTGGGAACAAAGTCGATTTTGTTACATCTAC

Clones of interest were then subcloned into pET29 expression vectors.

Alternatively, at the end of the last sort of a given round, the BD FACS Aria Cell Sorter was switched to plate mode, gates adjusted to only collect the top 0.1–0.3% of cells, and single cells collected in each well of a 96-well plate. After growing to saturation, these clones were subject to flow cytometry assays. Top performers were sequenced and then subcloned into pET29 expression vectors.

Reversion mutants

SrtAβ was subcloned into pET29 and used as PCR template for reactions with primers in Supplementary Table 3. Following USER assembly or KLD ligation, products were transformed into Thermo Fisher One-Shot Mach1 T1 Chemically Competent cells. Following sequence verification, the reversion mutants were amplified out of the pET29 backbone with HR primers and transformed into ICY200 with *Nhe*I/*Bam*HI-digested pCTCon2CTev vectors in a 5:1 insert:backbone mass ratio to yield yeast bearing single reversion mutants for flow cytometry analysis.

Minimal mutant

SrtA 4S.6 was subcloned into pET29 and used as PCR template for two reactions, one with primers ATCGTCCGAAC/ideoxyU/ATCAGTTTACCAACCTGCGCGCGGCGAAA AAAGGCAGC and AGGGTCAGC/ideoxyU/GTCTATCTTTGCCTTTCTGTTCATCCAGC ACTTCC, the other with primers AGCTGACCC/ideoxyU/GGCGACCTGCGATGATTAT

AACGTGGAAACCG and AGTTCGGACGA/ideoxyU/CAATCGCGGTATGGCCGATAA TGCTAATGTTCTGATCATCCAGGC. USER assembly of these two fragments yielded 4S.6 with S118I, G134R, K177R, and V182A mutations. This mutant version of 4S.6 was used as template for two further PCRs, one with primers ACCAGCATTTGTAA CG/ideoxyU/GAAACCGACCGCGGTGG and AAAAATTTTACTGGTT/ideoxyU/CCC ACACGCCGGTTTCCAC, the other with primers AAACCAGTAAAATTTT/ideoxyU/G TGGCGACCGAAGTGAAAGGATCC and ACGTTACAAATGCTGG/ideoxyU/CATTTTA TATTTACGGGTTTCGTTGC. USER assembly of these two fragments yielded the minimal mutant, 4S.6 with S118I, G134R, R159C, K177R, V182A, and R197S mutations. This mutant was then amplified out of the pET29 backbone with HR primers and transformed into ICY200 and ligated into *Nhe*I/*Bam*HI-digested pCTCon2CTev by homologous recombination.

Yeast transformation with LiAc/ss carrier DNA/PEG[56]

A 10 mL ICY200 starter culture in YPD (100 U/mL penicillin, 100 μg/mL streptomycin, and 50 μg/mL kanamycin) was grown overnight at 30 °C. Cells were centrifuged at 2500 g x 10 min, before removal of the supernatant and two washes with 25 mL water. Cells were resuspended in 1 mL water and transferred to a 1.5 mL Eppendorf tube. Cells were pelleted and washed once more before being resuspended in 1 mL of water and split into 100 μL aliquots. Aliquots were pelleted and supernatant removed. To each cell pellet was added 240 μL PEG 3550 (50% w/v), 36 μL LiOAc (1.0 M), 50 μL single stranded carrier DNA (2.0 mg/mL), 34 μL plasmid DNA or fragments (500–1000 ng) plus sterile water. Cells were then heat shocked at 42 °C for 40 min. Following heat shock, the cells were spun at 2500 g x 10 min, supernatant was removed, and the pellet was resuspended in 1 mL water. 10–100 μL of cell suspension was plated on SDC -Trp -Ura dropout plates and grown at 30 °C for 2–3 days.

Flow cytometry assays

Single clones were assayed by flow cytometry in a process similar to a library being prepared for sorting. Once a single clone was obtained via single cell sorting or lithium acetate transformation, it was grown to saturation in SDC -Ura -Trp dropout media and then induced overnight in SGR. Triglycine was conjugated to the cell surface by Sfp as with a library, with the volume scaled down depending on culture size. Reactions of surface-displayed sortases, TEV cleavage, and labeling are carried out as with a library preparation before analysis on a Bio-Rad ZE5 Cell Analyzer.

Sortase expression and purification

E. coli BL21(DE3) transformed with pET29 sortase expression plasmids were cultured at 37 °C in LB with 50 μg/mL kanamycin until OD_{600} = 0.5–0.8. IPTG was added to a final concentration of 1 mM and protein expression was induced overnight at 16 °C. The cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl supplemented with 1 mM $MgCl_2$, 2 units/mL DNAseI (NEB), 260 nM aprotinin, 1.2 μM leupeptin, and 1 mM PMSF). Cells were lysed by sonication and the clarified supernatant was purified on Ni-NTA agarose following the manufacturer's instructions. Fractions that were > 95% purity, as judged by SDS-PAGE, were consolidated and buffer

exchanged into 25 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM TCEP by size-exclusion chromatography in this buffer on a Superdex 200 Increase 10/300 GL column (GE). Enzyme concentrations were calculated by reducing agent-compatible BCA Protein Assay Kit (Pierce).

Aβ37-GGGRR cloning

Expression plasmid Aβ42/pET3 was amplified with primers GGRR-Fwd and GGRR-Rev. The PCR product was ligated with KLD enzyme mixture (New England BioLabs) and transformed into One-Shot Mach1 T1 Chemically Competent cells, from which Aβ37-GGGRR/pET3 was sequence verified and isolated.

GGRR-Fwd: CGCCGTTAATAGGAGCTCGATCCGG

GGRR-Rev: CCCACCGCCACCAACCATCA

Aβ expression and purification[44]

E. coli BL21(DE3) transformed with pET3 Aβ expression plasmids (AβM1–40, AβM1–42, or AβM1–37GGGRR) were cultured at 37 °C in LB-Carb until OD_{600} = 0.5–0.6. IPTG was added to a final concentration of 1 mM (AβM1–40 and AβM1–42) or 0.1 mM (AβM1–37GGGRR) and protein expression was induced for 4 hours at 37 °C. For AβM1–40 and AβM1–42, cells were pelleted and lysed by resuspension in 10mM Tris-HCl pH 8.0, 1 mM EDTA and sonication. Following lysis, the lysate was centrifuged for 10 minutes at 18,000 g. Supernatant was discarded and pellet was resuspended in 10mM Tris-HCl pH 8.0, 1 mM EDTA. Sonication, centrifugation, and removal of supernatant were repeated to yield an insoluble pellet. For AβM1–37GGGRR, cells were pelleted and lysed using B-PER bacterial protein extraction reagent (Thermo) supplemented with DNAseI and lysozyme and then centrifuged for 10 minutes at 18,000 g, with the insoluble pellet retained.

Insoluble pellets were resuspended in 8M urea, 10mM Tris/HCl pH 8.0, 1 mM EDTA and then sonicated. Solubilized inclusion bodies were diluted with 10mM Tris-HCl pH 8.0, 1 mM EDTA and added to pre-equilibrated DEAE-sepharose. After a 20–30 minute incubation, resin was batch filtered, washed for 5 minutes with 50 mM Tris pH 8.5, and then washed again for 5 minutes with 50 mM Tris pH 8.5, 25 mM NaCl. Following washes, recombinant peptides were eluted from resin with 50 mM Tris pH 8.5, 125 mM NaCl and lyophilized.

Chemically synthesized Aβ

Aβ1–40 and Aβ1–42 peptides (including Btn-LC-Aβ40 and Btn-LC-Aβ42) were synthesized and purified using reverse-phase HPLC by Dr. James I. Elliott at the ERI Amyloid laboratory, Oxford, CT, USA. Peptide mass and purity (>99%) were confirmed by reverse-phase HPLC and electrospray ion trap mass spectrometry.

Isolation of Aβ monomers[28]

Lyophilized Aβ peptides, whether synthetic in origin (ERI Amyloid Laboratory, LLC), or produced by recombinant technology, were dissolved in 7 M guanidium chloride, 50 mM Tris pH 7.5, 2 mM EDTA at a concentration of 1 mg/mL and incubated overnight.

Denatured Aβ was then purified by size exclusion chromatography using a Superdex 75 300/10 column (GE) at a flow rate of 0.5 mL/min in alkaline buffer (50 mM Tris-HCl pH 8.5) to minimize peptide aggregation. Peptide concentration was measured by A_{275} (ε = 1361 $M^{-1}cm^{-1}$). Peptide was either used immediately after purification or diluted to 20 μM, aliquoted, and frozen at −80 °C for later use.

Western blot analysis for fetuin A

Samples of sortase reactions with fetuin A were added to 4x NuPAGE lithium dodecyl sulfate (LDS) buffer (Invitrogen, NP0007), heat denatured, and loaded onto a 4–12% bis-tris gel and ran at 160 V for 30 min in MES running buffer. Samples from reactions in plasma were diluted as follows: 20 μL sample diluted + 30 μL TBS + 20 μL LDS buffer, for a total dilution of 3.5x. Gels were transferred to PVDF membrane via iBlot and membrane blocked with Superblock Blocking Buffer (Thermofisher, 37515) for 1 hour at room temperature. The membrane was then incubated with mouse anti-fetuin A antibody (Abcam, ab89227, 1:500 dilution in Superblock TBS + 0.1% tween-20) overnight at 4°C followed by washing in PBS-T (PBS + 0.1% tween-20) three times for 5 minutes each. Secondary antibodies Streptavidin-IR800 (Licor, 926–32230) and goat anti-mouse-IR680LT (Licor, 926–68020) (both 1:10,000 dilution in Odyssey Block in PBS (Licor, 927–40000), 0.1% Tween-20, 0.01% SDS) were applied for 30 min at room temperature in the dark. The membrane was washed with PBS-T three times for 5 minutes each, followed by one wash with MilliQ water and imaged on an Odyssey Imager.

Western blot analysis for Aβ

Samples of sortase reactions with Aβ were added to 4x LDS buffer and, without heat denaturing, loaded onto a 4–12% bis-tris gel and ran at 160 V for 30 min in MES running buffer. Gels were transferred to PVDF membrane via iBlot and membrane blocked with Superblock Blocking Buffer (Thermofisher, 37515) for 1 h at rt. The membrane was then incubated with mouse anti-Aβ 4G8 antibody (Biolegend, 800702, 1:1000 dilution in Superblock TBS + 0.1% tween-20) overnight at 4°C followed by washing in PBS-T (PBS + 0.1% tween-20) three times for 5 minutes each. Secondary antibodies Streptavidin-IR800 (Licor, 926–32230) and goat anti-mouse-IR680LT (Licor, 926–68020) (both 1:10,000 dilution in Odyssey Block in PBS (Licor, 927–40000), 0.1% Tween-20, 0.01% SDS) were applied for 30 min at room temperature in the dark. The membrane was washed with PBST three times for 5 minutes each, followed by one wash with MilliQ water and imaged on an Odyssey Imager.

Streptavidin pulldown of sortase-labeled plasma proteins

1 mL of normal human plasma was combined with 10 μL 1M $CaCl_2$, 10 μL of 0.1M GGGK(Biotin), and 10 μL of 100 μM sortase 4S.6 or SrtAβ, then incubated at room temperature for 2 hours. 100 μL of pre-equilibrated Ni-NTA resin slurry was added to the mixture and incubated at room temperature with shaking for 15 minutes before being filtered through a 0.2 μm spin filter before dilution to 10 mL final volume in PBS-E (PBS + 1 mM EDTA). The solution was concentrated using a 3kDa molecular weight cut-off spin concentrator for 30 minutes at 3500xg and a final volume of <1 mL. This sample was diluted with PBS-E to 10 mL final volume, re-concentrated, and re-diluted in a total of six

wash cycles to give an expected small molecule biotin concentration of <1 nM. The concentrated mixture was then incubated with 200 μL of pre-equilibrated Invitrogen MyOne Streptavidin C1 Dynabeads with shaking for 30 minutes before magnetic separation and washing three times with PBS + 0.1% Tween-20. Beads were then resuspended in 100 μL SDS-PAGE loading buffer with 100 μM free biotin and incubated at 95 °C for 15 minutes. A 15-μL aliquot was then run on a 4–12% Bis-Tris PAGE gel and visualized by staining with coomassie blue.

HPLC assay of sortases on LMVGG

Reactions were performed with Abz-LMVGGK(Dnp)-$CONH_2$ peptide (custom synthesis from Genscript) fixed at 10 μM. Reaction conditions were 300 mM Tris pH 7.5, 150 mM NaCl, 100 mM H_2N-GGG-COOH, 5 mM $CaCl_2$, 5% v/v DMSO. 5 μL of 10 μM sortase stock was added to 45 μL reaction buffer, yielding a final enzyme concentration of 1 μM. Reactions were incubated for 120 min at 22.5 °C. Reactions were quenched with 10 μL 1 N HCl. The total volume of each reaction was transferred to HPLC sample vials and ran on analytical reverse phase Agilent Zorax SB-C18 (2.1 × 150 mm, 5 μm) and chromatographed using a linear gradient 10 to 56.5% acetontrile with 0.1% TFA in 0.1% aqueous TFA over 13 minutes. To calculate the percent conversion, the ratio of the integrated areas of the GK(Dnp)-$CONH_2$ (rt = 6.7 minutes) and Abz-LMVGG(Dnp)-$CONH_2$ (rt = 11.6 minutes) Abs_{355} peaks were compared directly.

HPLC assay of sortases on Aβ40

Reactions were performed with 20 μM Aβ40 and 1 mM GGGK(Dnp) in 50 mM Tris pH 8.5, 150 mM NaCl, and 5 mM $CaCl_2$. 5 μM of SrtAβ was added to this mixture and incubated at room temperature overnight. Reactions were quenched with 10 μL 1 N HCl. The total volume of each reaction was transferred to HPLC sample vials and ran on analytical reverse phase Agilent Zorax SB-C18 (2.1 × 150 mm, 5 μm) and chromatographed using a linear gradient 10 to 56.5% acetontrile with 0.1% TFA in 0.1% aqueous TFA over 13 minutes. To calculate the percent conversion, the ratio of the integrated areas of the GGGK(Dnp) (rt = 8.2 minutes) and Aβ37-GGGK(Dnp) (rt = 12.6 minutes) Abs_{355} peaks were compared directly.

Kinetic assay of sortases on LPESG

Assays to determine k_{cat} and $K_{m\ LPESG}$ were performed in 300 mM Tris pH 7.5, 150 mM NaCl, 5 mM $CaCl_2$, 5% v/v DMSO, and 10 mM Gly-Gly-Gly-COOH (GGG). The concentration of the LPESG peptide substrate ranged from 62.5 μM to 4 mM, and enzyme concentrations ranged from 100 nM to 1000 nM. Reactions were initiated with the addition of enzyme and incubated at 22.5 °C for 7 minutes (sortase 4S.6) or 2 hours (SrtAβ) before quenching with 0.2 volumes of 5 M HCl. 5 to 10 nmol of peptide from the quenched reactions were injected onto an analytical reverse-phase Eclipse XDB-C18 HPLC column (4.6×150 mm, 5 μm, Agilent Technologies) and chromatographed using a linear gradient of 10 to 65% acetonitrile with 0.1% TFA in 0.1% aqueous TFA over 13 minutes. Retention times under these conditions for the Abz-LPESGK(Dnp)-$CONH_2$ substrate and the released GKDnp peptide were 12.8 and 10.4 minutes, respectively. To calculate the percent conversion, the ratio of the integrated areas of the GK(Dnp)-$CONH_2$ and Abz-

LPESGK(Dnp)-$CONH_2$ peptide Abs_{355} peaks were compared directly. To determine k_{cat} and $K_{m,LPESG}$, reaction rates were fit to the Michaelis-Menten equation in GraphPad Prism.

Kinetic assay of sortases on LMVGG

Assays to determine k_{cat} and $K_{m\ LMVGG}$ were performed in 300 mM Tris pH 7.5, 150 mM NaCl, 5 mM $CaCl_2$, 5% v/v DMSO, and 10 mM Gly-Gly-Gly-COOH (GGG). The concentration of the Abz-LMVGG(Dnp)-$CONH_2$ peptide substrate ranged from 10 to 200 μM with enzyme concentration of 1 μM. Reactions were conducted in 96-well half area black/clear flat bottom plates (Corning) and initiated with the addition of enzyme. Plates were incubated at 24 °C and monitored for increase in fluorescence (ex=317 nm, em=420 nm) in a Tecan plate reader for 2 hours. Changes in fluorescence were converted to molar velocities using calibration curves of Abz-LMVGG(Dnp)-$CONH_2$ and a 1:1 mixture of free Abz and Dnp. Inner filter quenching effects were corrected using $F_{corr} = F_{obs}$ x antilog[(A_{ex} + A_{em})/2], where F_{corr} is the corrected fluorescence value, F_{obs} is the observed fluorescence value, A_{ex} is the absorbance at 317 nm, and A_{abs} is the absorbance at 420 nm. To determine k_{cat} and $K_{m,LMVGG}$, initial velocities were fit to the Michaelis-Menten equation in GraphPad Prism.

Semi-synthesis of Aβ(M1–37- GGGK(Btn))

Freshly purified Aβ(M1–40) monomers (120 μM in 50 mM Tris pH 8.5) were supplemented with 150 mM NaCl, 5 mM $CaCl_2$, and 1 mM TCEP and reacted overnight at room temperature with 50μM SrtAβ and 1 mM GGGK(Btn). After desalting in a 3 kDa molecular weight cutoff spin filter, the reaction mixture was lyophilized and then dissolved in 7 M guanidium chloride, 50 mM Tris pH 7.5, 2 mM EDTA and ran on a Kinetex C18 100 Å (150×30 mm, 5 μm, Phenomenex) column. The acetonitrile concentration was increased from 10 to 35% over the first 5 minutes, 35 to 38% over the next 6 minutes, and then from 38 to 90% over the next 5 minutes. The major peak eluted at 12.8 minutes. This was confirmed to be Aβ(M1–37- GGGK(Btn)) by LC/MS (*m/z* = 4731.98 observed, 4730.27 expected) and the product was lyophilized and stored at −20 °C for later use.

Streptavidin capture ELISA for detection of biotinylated Aβ

Aβ(M1–37- GGGK(Btn)) standards were prepared in diluent (TBS+0.1% Tween-20+1% BSA) in a range of concentrations from 20 nM to 312 pM. Samples were diluted as necessary in this same diluent. Pre-blocked Streptavidin Coated High Capacity plates (clear, 96-well, Pierce) were washed 2x with TBS-T (TBS + 0.1% Tween-20) before addition of standards and any samples. Biotinylated material was captured by streptavidin at room temperature for 2 hours. Plates were washed three times with TBS-T. 100 μL of mouse anti-Aβ clone 4G8 (1:2000 in diluent, Biolegend, 800702) was added to each well and incubated at room temperature for one hour. Following 3x TBS-T washes, each well was treated with 100 μL of goat anti-mouse IgG HRP conjugate (1:4000 in diluent, ThermoFisher, A-10668) for 30 minutes at room temperature. Plates were washed four times with TBS-T before addition of 50 μL TMB (ThermoFisher, 34028). Wells were allowed to develop until saturation and then quenched with 50 μL of 2M H_2SO_4. The absorbance of each well at 450 nm was then measured using a Tecan Plate Reader. The standard curve was fitted to 4-

parameter logistics curve by Solver in Excel and used to calculate concentration of biotinylated Aβ in present samples.

m266 antibody capture ELISA for detection of biotinylated Aβ

Thermo Nunc Maxisorp plates (96-well, clear) were incubated overnight with 100 μL of 3 μg/mL anti-Aβ antibody m266. The next day, plates were washed 3x with TBS-T and blocked for 2 hours with 5% MSD Blocker A (Meso Scale, Rockville, MD) in TBS-T. Aβ(M1–37- GGGK(Btn)) standards were prepared in diluent (1% MSD Blocker A in TBS-T) in a range of concentrations from 2.5 ng/mL to 39 pg/mL. Samples were diluted as necessary in this same diluent. Plates were washed 3x with TBS-T before addition of standards, samples, and blanks in triplicate. After 2 hours of capture, plates were washed 3x with TBS-T. 100 μL of streptavidin-HRP (1:100 in diluent, R&D Systems Part # 890803) was added to each well for 30 minutes. After 4x TBS-T washes, wells were developed with 50 μL TMB (Thermo N301) and quenched with 2M H_2SO_4. The absorbance at 450 nm of each well was then measured by a Molecular Devices plate reader. The standard curve was fitted to 4-parameter logistics curve by Solver in Excel and used to calculate the concentration of biotinylated Aβ present in samples.

ELISA assays for Aβ40 and Aβ42[40]

Thermo Nunc Maxisorp plates (96-well, clear) were incubated overnight with 100 μL of 3 μg/mL anti-Aβ antibody m266. The next day, plates were washed 3x with TBS-T and blocked for 2 hours with 5% MSD Blocker A in TBS-T. Aβ40 and Aβ42 standards were prepared in diluent (1% MSD Blocker A in TBS-T) in a range of concentrations from 2.5 ng/mL to 39 pg/mL. Samples were diluted as necessary in this same diluent. Plates were washed 3x with TBS-T before addition of standards, samples, and blanks in triplicate. After a 2-hour capture, plates were washed and secondary antibodies (1:2500 biotinylated 21F12 for Aβ42, 1:4000 biotinylated 2G3 for Aβ40) were added for 2 hours. After another set of washes, streptavidin-HRP (1:100) was added for 30 minutes. Plates were then washed, developed with TMB, and quenched with H_2SO_4. The absorbance at 450 nm of each well was measured by a Molecular Devices plate reader. The standard curves were fitted to 4-parameter logistics curve by Solver in Excel and used to calculate concentrations of Aβ40 and Aβ42 present in the samples.

CSF and plasma samples

Human CSF specimens were obtained in accordance with local clinical regulations approved by the Partners Institutional Review Board (Walsh, BWH2017P0000259). All samples were from the Biobank at Partners HealthCare in Boston, Massachusetts. Donors had no history of diseases of the central nervous system. Donor demographic information can be found in Supplementary Table 4. Pooled human plasma was purchased from GeneTex (Cat No. GTX73265).

CSF labeling with GGG using SrtAβ

Aliquots of CSF collected from 10 different patients were supplemented with 5 mM $CaCl_2$ and treated for 1 hour with 5 μM SrtAβ and 500 μM GGG. Reactions were quenched with

addition of 5 mM EDTA and diluted 2-fold with 1% MSD Blocker A in TBS-T. Part of the sample was set aside for Aβ42 measurement, while the rest was diluted 5-fold (total dilution = 10-fold) for Aβ40 measurement. Untreated aliquots from the same patients were diluted similarly. Aβ40 and Aβ42 were captured by anti-Aβ antibody m266 and detected with C-terminal specific antibodies as described above.

CSF labeling with GGGK(Btn) using SrtAβ

Aliquots of CSF collected from 10 different patients were supplemented with 5 mM $CaCl_2$ and treated for 2 hours with 5 μM SrtAβ and 500 μM GGGK(Btn). Reactions were quenched with addition of 5 mM EDTA and all samples (full reactions, no SrtAβ control, no GGGK(Btn) control, and untreated) were diluted 10-fold with 1% MSD Blocker A in TBS-T. Biotinylated Aβ was captured by anti-Aβ antibody m266 and detected without secondary antibody using streptavidin-HRP as described above.

Semi-synthesis of Aβ(M1–37-GGGRR)

Immediately following elution from DEAE resin in 50 mM Tris pH 8.5 + 125 mM NaCl, recombinant Aβ42 (20 mL of estimated concentration 40 μM = 3–4 mg) was supplemented with 5 mM $CaCl_2$ and 5 mM DTT and treated overnight at room temperature with 20 μM SrtAβ and 200 μM GGGRR. The reaction mixture was concentrated to 1 mL in a 3 kDa molecular weight cutoff spin concentrator, diluted to 20 mL with milliQ water to reduce the salt concentration, and then concentrated back to 1 mL and lyophilized. The lyophilized reaction mixture was then denatured overnight in 7 M guanidium chloride, 50 mM Tris pH 7.5, 2 mM EDTA and ran on a Zorbax 300SB-C18 (9.4×250 mm, 5 μm, Agilent) column. After 5 minutes at 10% acetontrile with 0.1% TFA in 0.1% aqueous TFA, the acetonitrile concentration was increased to 30% over 5 minutes, and then to 50% over 20 minutes. Aβ(M1–37-GGGRR) eluted at 17.5 minutes. Aβ(M1–37-GGGRR) identity was confirmed by LC/MS (m/z = 4689.85 observed, 4688.30 expected) and the fraction containing it was lyophilized and stored at −20 °C for later use..

ThT Assay[45]

Aβ peptides were denatured and SEC-isolated in 20 mM sodium phosphate pH 8.0. Concentrations were determined by A_{275} and stock solutions of 20.2 μM peptide in elution buffer were prepared. To 990 μL of each stock solution was added 10 μL of thioflavin T (2 mM in water), yielding 1 mL of 20 μM peptide and 20 μM ThT. 20 μM ThT in elution buffer was used as diluent to make 10 μM peptide samples. Samples were aliquoted 120 μL per well to a sterile Nunc 96 well black polystyrene plate (Thermo Scientific, Cat. # 237105). A Molecular Devices plate reader was used to follow change in fluorescence (435 ex/480 em) over 48–60 hours.

Negative contrast transmission electron microscopy of Aβ fibrils[57]

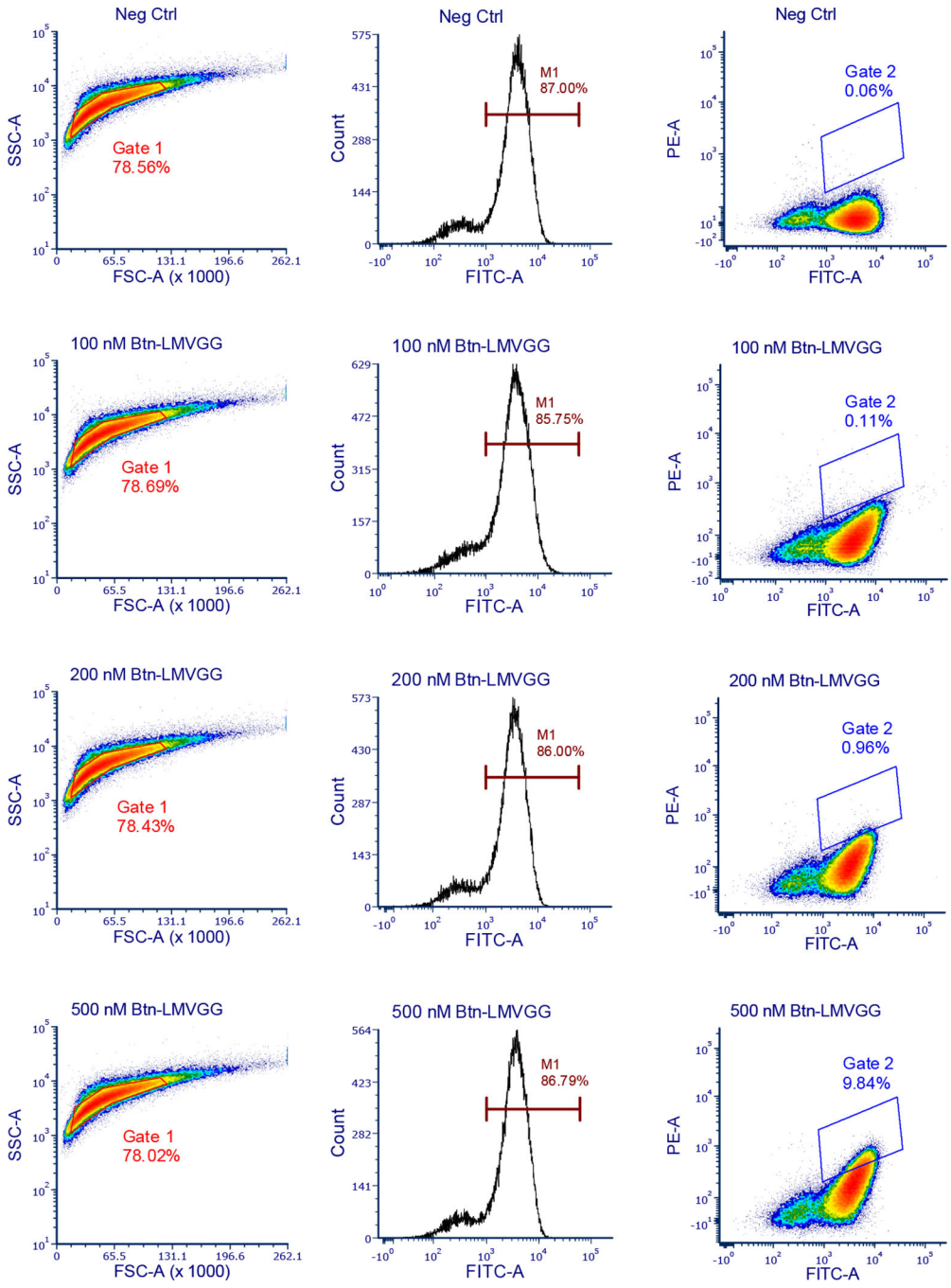
Samples of Aβ(M1–37-GGGRR) (n=6) and Aβ(M1–42) (n=2) lacking ThT were included alongside ThT containing samples in the assay described above. Following aggregation, these samples were applied to carbon-coated Formvar grids, left for 1 minute, fixed with glutaraldehyde, washed with MQ water, and wicked dry with filter paper. 2% uranyl acetate

was then added and incubated for two minutes. The grid was wicked dry and allowed to air-dry for 10 minutes. Grids were stored in a sealed container and viewed under a Tecnai G2 BIOTWIN electron transmission microscope operated at 80 kV. All reagents were supplied by Electron Microscopy Sciences (Hatfield, PA).

Data Availability Statement

Source data has been provided for Figures 2–4, as well as Extended Data Figures 2, 3, and 5–9. Additional data (e.g. Sanger sequencing relevant to Table 1 or .fcs files from the evolution) is available upon request.

Extended Data

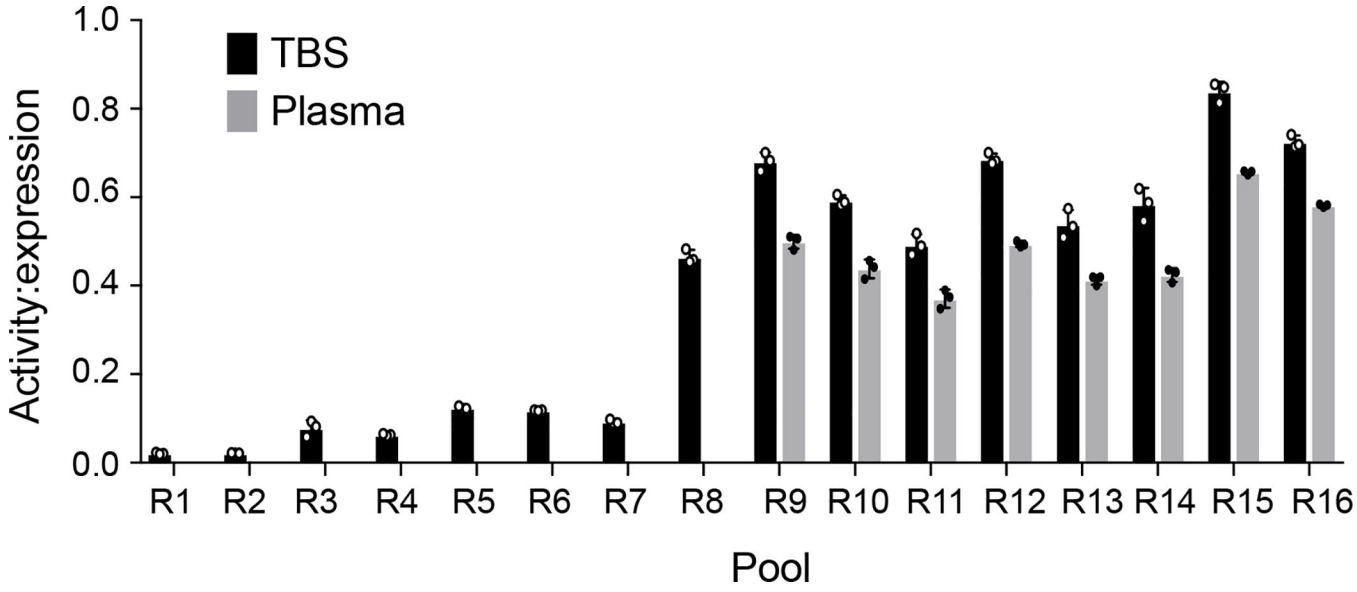


Extended Data Fig. 1: Gating strategy for SrtAβ evolution.

Following loading of cell surfaces with triglycine, the library is split into several aliquots that are treated with varying amounts of positive selection substrate. Surface-displayed sortase reactions, TEV removal or sortase, and staining are conducted in parallel on these aliquots. The aliquots are then analyzed before sorting. Populations are gated first on the

basis of size (FSC vs. SSC), and then successful induction is confirmed on a FITC histogram.

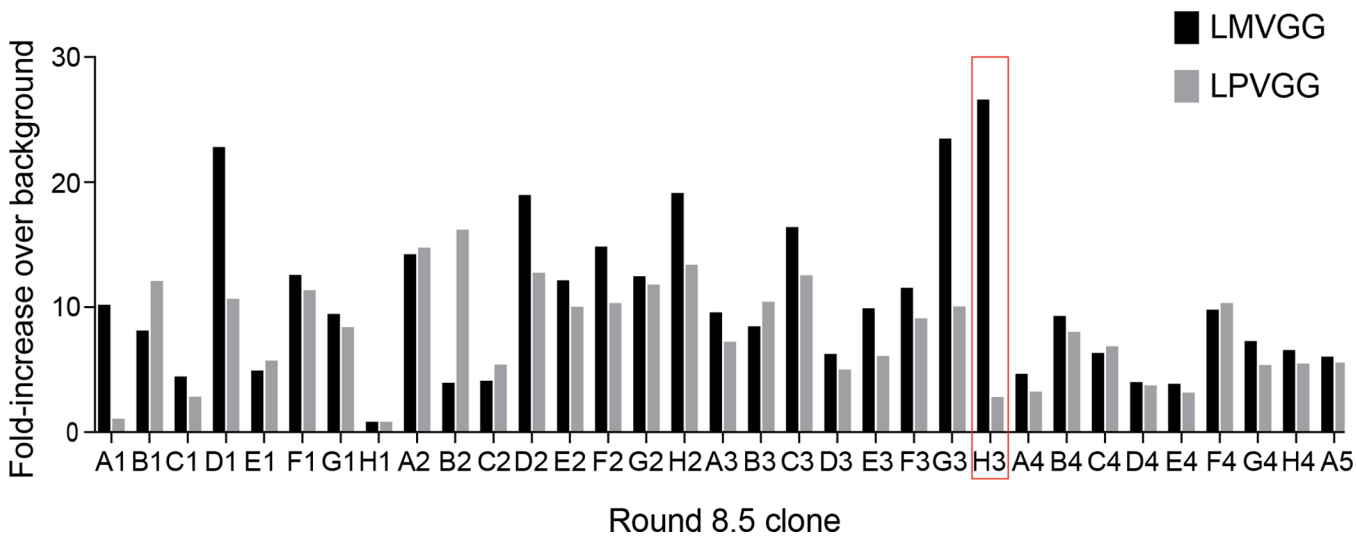
FITC vs PE dot plots are made for each aliquot, and a polygon gate is drawn that includes <0.1% of the negative control (no positive selection substrate) aliquot while roughly matching its slope. This gate is then applied to the other aliquots. The aliquot with the least amount of positive selection substrate that shows >10-fold increase in gated events over the negative control is used for sorting. In the case of this representative data from round 11, that would be the 200 nM Btn-LMVGG aliquot.



Extended Data Fig. 2: Flow cytometry analysis of library activity on LMVGG over the course of evolution.

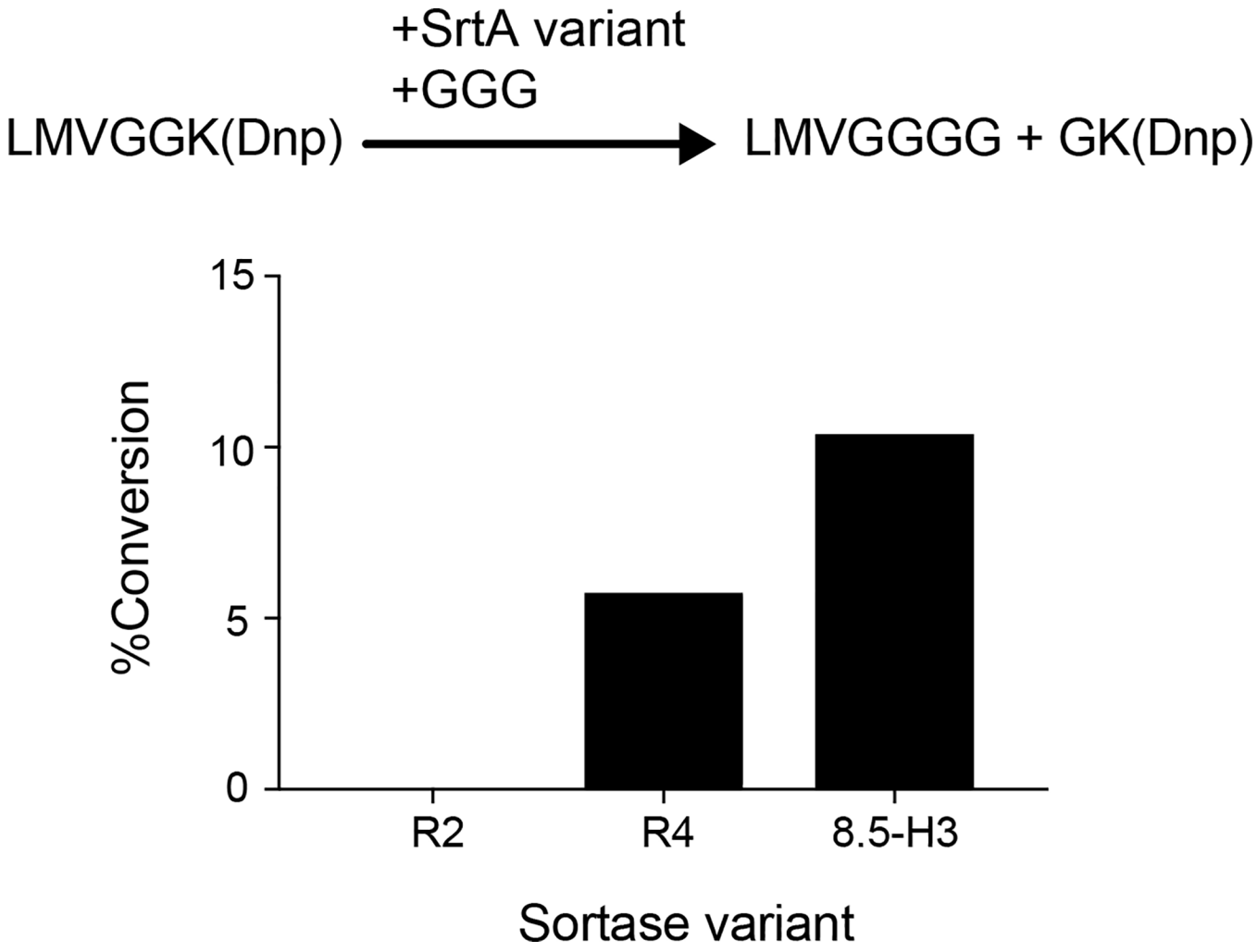
Stocks of yeast libraries remaining at the end of each round (R) were regrown, induced, and subjected to cell surface sortase reactions.

Each library was incubated for 1 hour with 1 μM Btn-LMVGG. All libraries were assayed for LMVGG activity in TBS, while the libraries that survived negative selection against fetuin A and other plasma proteins (R9–16) were also assayed in human plasma. Data shown are the mean ± standard deviation of three replicates. Activity:expression is the transpeptidation activity (PE) to sortase expression (FITC) ratio. Overall, the trend towards improved LMVGG activity per unit expression is strong, indicative of a successful evolution campaign.



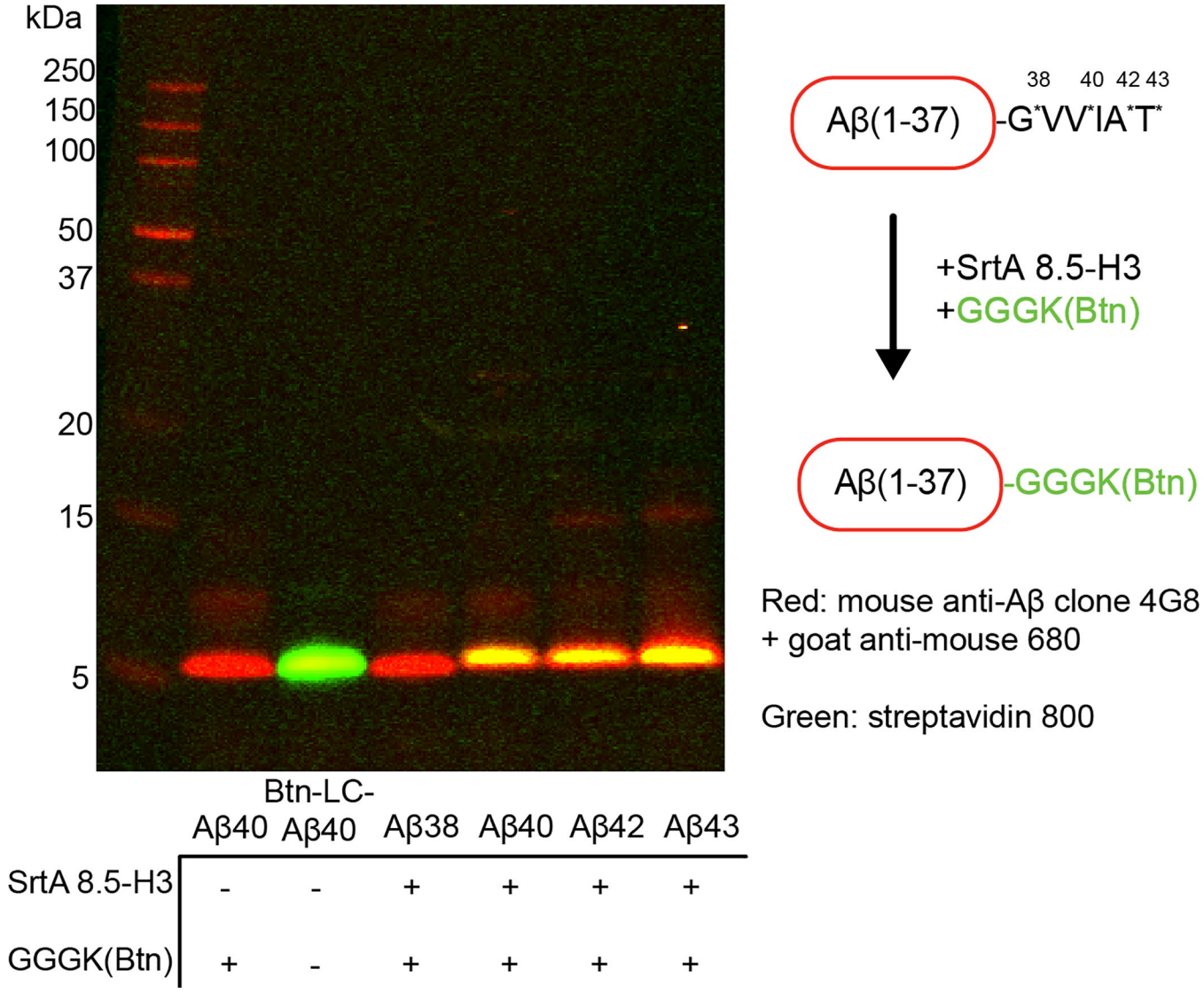
Extended Data Fig. 3: Flow cytometry analysis of round 8 single clones.

At the end of round 8 sort 5 (8.5), single cells were sorted into a 96-well plate, grown to saturation, and induced. After attachment of GGGK-CoA via Sfp ligation, the clonal populations were given an hour to react with 500 nM Btn-LMVGG or 500 nM Btn-LPVGG. Flow cytometry analysis revealed that clone 8.5-H3 possessed the best combination of high activity on the LMVGG substrate with low activity on the LPVGG substrate. Activity is defined as fold-increase in PE signal over a negative control (0 nM Btn-LMVGG) aliquot of each variant.



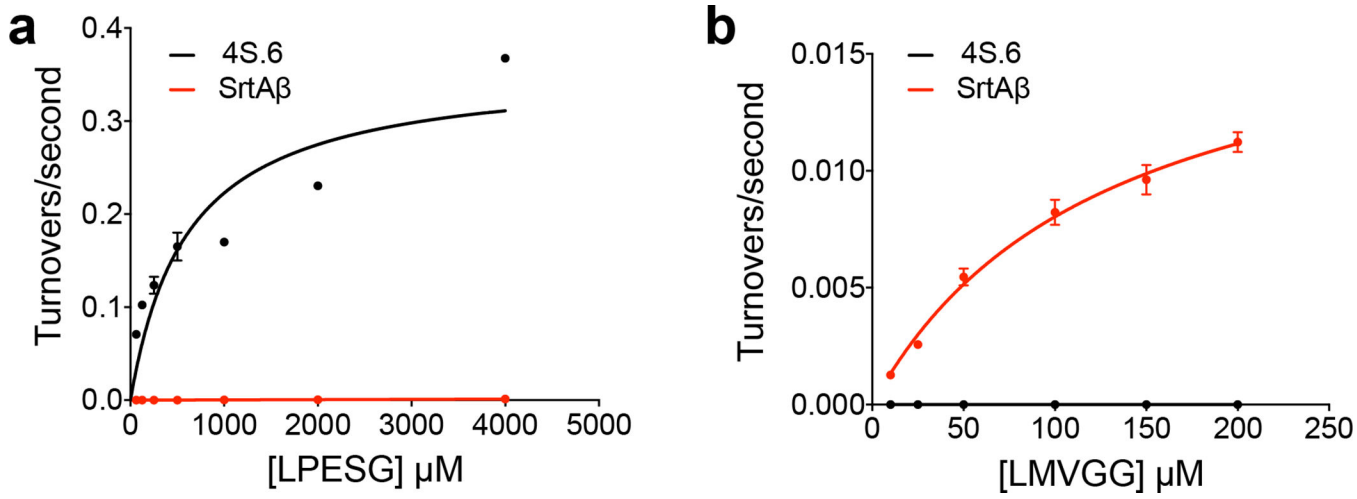
Extended Data Fig. 4: HPLC assay of LMVGG activity.

In this assay, sortase-mediated transpeptidation of a chromophore-linked substrate (Abz-LMVGGK-Dnp) liberates free GK-Dnp, the formation of which can be followed by HPLC[1]. The round 2 consensus sequence (R2, 4S.6 + R94Y, S118I, A122W, D124G, G134R, and V189F mutations) did not show detectable activity in this assay. The first variants with detectable activity on LMVGG in this assay emerged from round 4; the most active of these variants (R4) displayed modest activity on Abz-LMVGGK-Dnp, with 1 μM of enzyme converting 5.8% of 10 μM substrate to product in two hours. Notably, this variant contained the V182A, T196S, and R197S mutations. By round 8, activity in this assay roughly doubled, with 1 μM of clone 8.5-H3 converting 10.4% of 10 μM substrate to product in two hours.



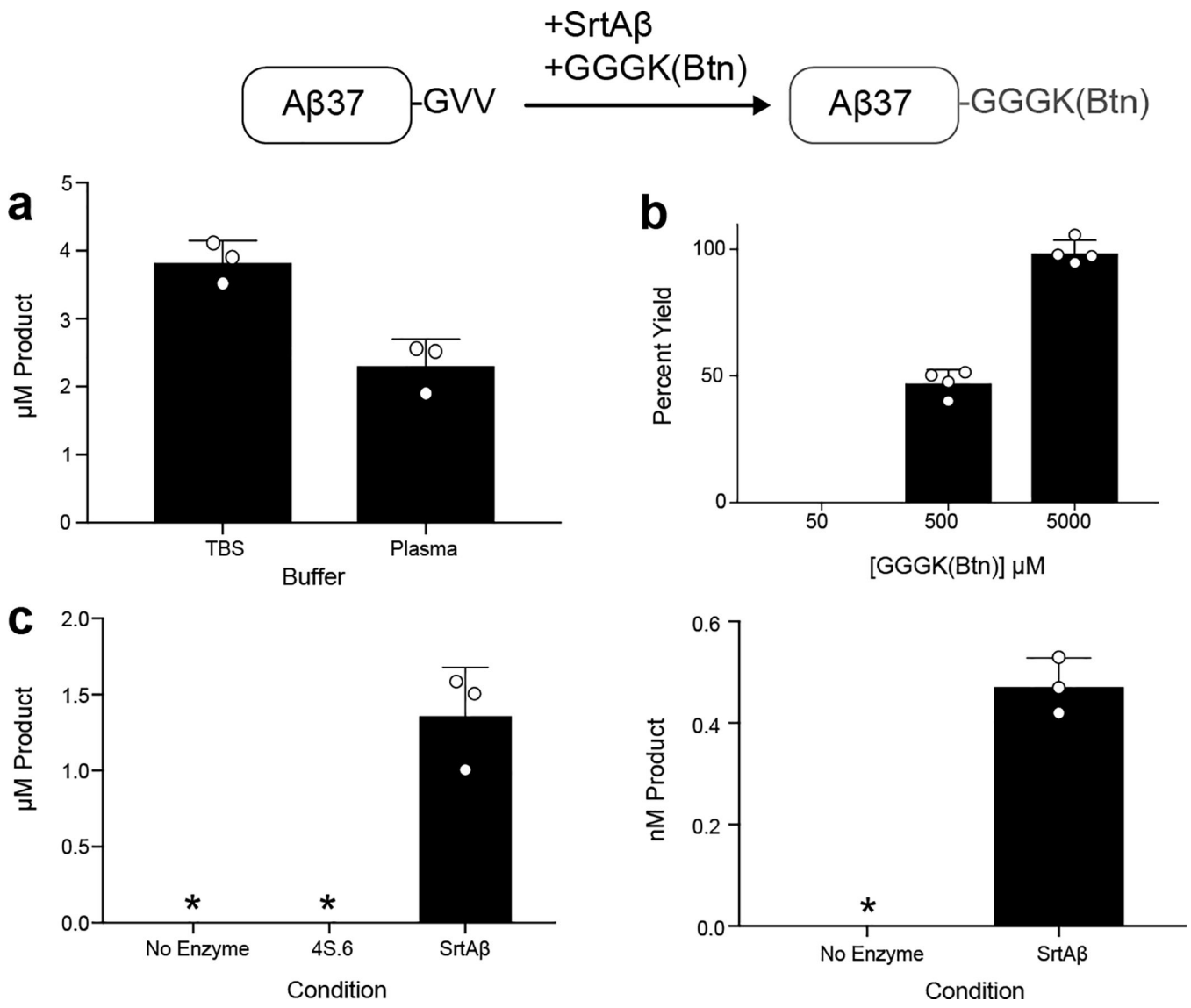
Extended Data Fig. 5: Evolved sortase labels Aβ alloforms with C-termini that extend to Val40 and beyond.

SrtA 8.5-H3 (10 μM) was incubated with GGGK(Btn) (100 μM) and different alloforms of Aβ (Aβ38, Aβ40, Aβ42, or Aβ43, 10 μM each) for 2 hours in DPBS + 2 mM CaCl2. Labeling was observed for Aβ40, Aβ42, and Aβ43. Aβ38 is not labeled, which is expected given that the final glycine of the LMVGG recognition sequence is not amidated in this alloform. This experiment was repeated three times independently with similar results.



Extended Data Fig. 6: Kinetic analysis of sortase 4S.6 and SrtAβ.

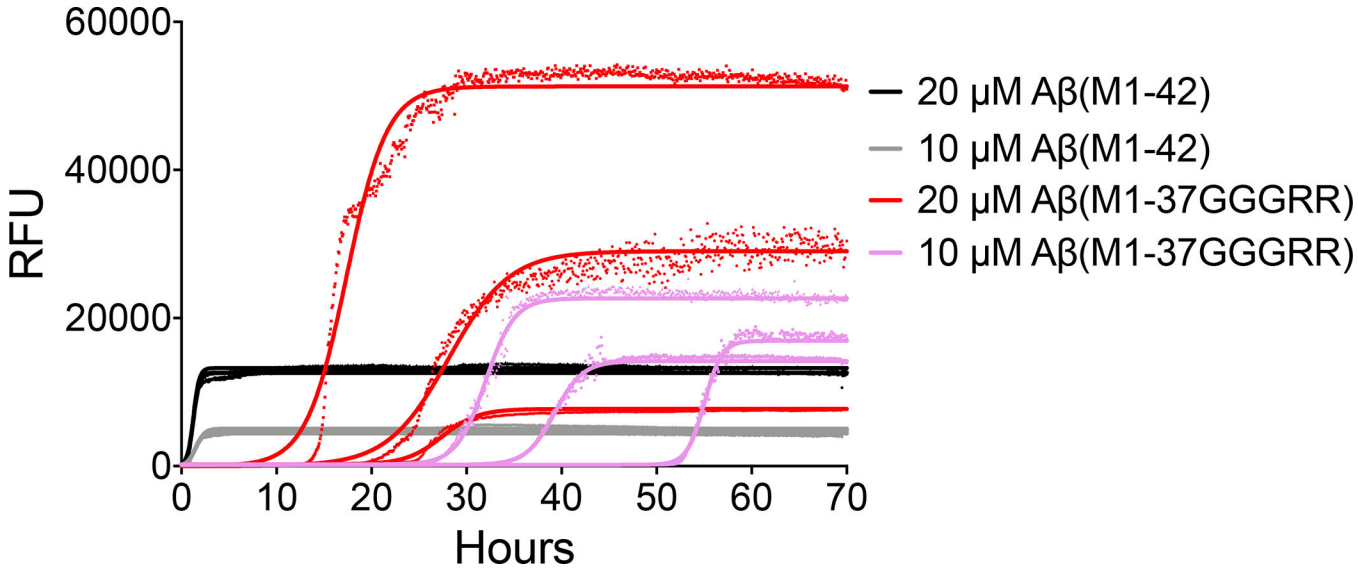
(**a**) Sortase 4S.6 shows kcat = 0.36 s^{-1} (95% CI = 0.22 to 0.96 s^{-1}) and KM = 610 μM (95% CI = 90 to 5550 μM) on its cognate substrate LPESG in an established HPLC assay for sortase transpeptidation activity. SrtAβ activity was detectable in this assay, but at levels too low to accurately determine its kinetic parameters on LPESG. Attempts to establish kinetic parameters for LMVGG in the HPLC assay were complicated by this substrate's limited solubility in the reaction buffer. (**b**) A more sensitive fluorescence method was used to detect smaller amounts of turnover at lower substrate concentrations. Importantly, the solubility of the LMVGG substrate is within a range where inner filter quenching effects can be corrected. Using this method, SrtAβ shows kcat = 0.018 s^{-1} (95% CI = 0.015 to 0.023 s^{-1}) and KM = 128 μM (95% CI = 87 to 198 μM) on LMVGG. Sortase 4S.6 activity on LMVGG was not detectable in this assay. Points and error bars represent the average of three replicates ± one standard deviation.



Extended Data Fig. 7: Aβ40 labeling in human plasma.

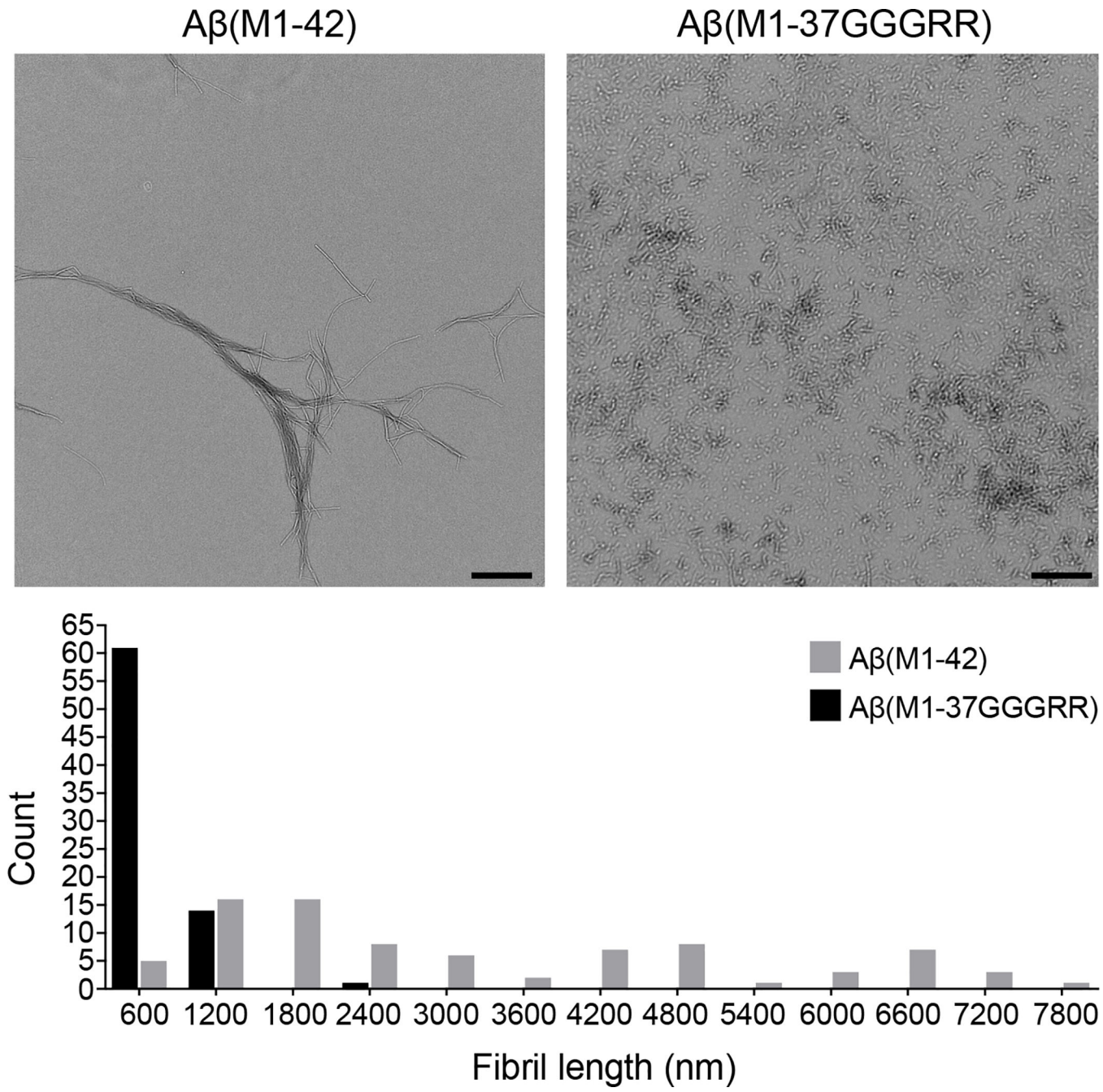
(**a**) Aβ40 (5 μM) added to TBS or plasma was labeled with GGGK(Btn) (50 μM) by SrtAβ (1.5 μM) for two hours. Product was captured on a streptavidin coated plate and detected using the anti-Aβ antibody 4G8. The reaction was 1.6-fold more efficient in TBS than in plasma.

(**b**) Aβ40 (5 μM) added to human plasma was labeled in the presence of various concentrations of GGGK(Btn) by an evolved sortase variant from round 14 (5 μM). After 2 hours, biotinylated product was undetectable in the reaction with a 10-fold excess of GGGK(Btn), whereas almost all of the Aβ40 was biotinylated when a 1000-fold excess of GGGK(Btn) was used. (**c**) Aβ40 (5 μM or 5 nM) added to plasma was labeled with GGGK(Btn) (500 μM) by SrtAβ (1 μM) for two hours. No product was detected in the absence of enzyme or when using the starting enzyme 4S.6 in place of SrtAβ (* indicates below limit of detection). Each column represents the average of at least three replicates. Error bars represent one standard deviation.



Extended Data Fig. 8: Aggregation of AβM1–42 compared to AβM1– 37GGGRR.

The aggregation of recombinant Aβ42 and recombinant Aβ37GGGRR monomers was monitored by thioflavin T binding. Curves were fitted to each replicate (n =3 for both peptides at both concentrations) by the Boltzmann equation. The aggregation of recombinant Aβ37GGGRR monomers is retarded relative to recombinant Aβ42, with an average t1/2 = 24.3 hours at 20 μM and 38.4 hours at 10 μM (compared to 1.25 hours and 1.46 hours for recombinant Aβ42 at 20 μM and 10 μM, respectively). These results are consistent with those observed for semi-synthetic Aβ37GGGRR created by using SrtAβ to label recombinant Aβ42 with GGGRR.



Extended Data Fig. 9: Electron microscopy of amyloid fibrils.

End-point samples from the ThT time-course study of AβM1–42 and AβM1–37GGGRR show marked differences in ultra-structure. Scale bars represent 800 nm. The most apparent difference between the two sets of fibrils is size. 61/76 AβM1–37GGGRR structures measured were less than 600 nm long, with an average length of 380 nm. In contrast, the average length of AβM1–42 fibrils was ~7 times larger at 2670 nm (n = 83). Aggregation and electron microscopy of amyloid fibrils was performed twice with similar results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements.

We thank Dr. Tiernan O'Malley and Dr. Greg Newby for helpful comments. We thank Anahita Vieira for assistance with editing this manuscript. This work was supported by R01EB022376, and R35GM118062, R01AG046275 and HHMI. DMW is an Alzheimer's Association Zenith Fellow.

References Cited

1. Boutureira O & Bernardes GJ Advances in chemical protein modification. Chem Rev 115, 2174–2195, doi:10.1021/cr500399p (2015). [PubMed: 25700113]
2. Chin JW Expanding and reprogramming the genetic code of cells and animals. Annu Rev Biochem 83, 379–408, doi:10.1146/annurev-biochem-060713-035737 (2014). [PubMed: 24555827]
3. Noren CJ, Anthony-Cahill SJ, Griffith MC & Schultz PG A general method for site-specific incorporation of unnatural amino acids into proteins. Science 244, 182–188, doi:10.1126/science.2649980 (1989). [PubMed: 2649980]
4. Shah NH & Muir TW Inteins: Nature's Gift to Protein Chemists. Chem Sci 5, 446–461, doi:10.1039/c3sc52951g (2014). [PubMed: 24634716]
5. Zhang C et al. Π-Clamp-mediated cysteine conjugation. Nat Chem 8, 120–128, doi:10.1038/nchem.2413 (2016). [PubMed: 26791894]
6. Distefano MD Recent progress in enzymatic labelling techniques and their applications. Chem Soc Rev (2018).
7. Paterson GK & Mitchell TJ The biology of Gram-positive sortase enzymes. Trends Microbiol 12, 89–95, doi:10.1016/j.tim.2003.12.007 (2004). [PubMed: 15036325]
8. Mazmanian SK, Liu G, Ton-That H & Schneewind O Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. Science 285, 760–763, doi:10.1126/science.285.5428.760 (1999). [PubMed: 10427003]
9. Ton-That H, Liu G, Mazmanian SK, Faull KF & Schneewind O Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of Staphylococcus aureus at the LPXTG motif. Proc Natl Acad Sci U S A 96, 12424–12429 (1999). [PubMed: 10535938]
10. Kruger RG et al. Analysis of the substrate specificity of the Staphylococcus aureus sortase transpeptidase SrtA. Biochemistry 43, 1541–1551, doi:10.1021/bi035920j (2004). [PubMed: 14769030]
11. Glasgow JE, Salit ML & Cochran JR In Vivo Site-Specific Protein Tagging with Diverse Amines Using an Engineered Sortase Variant. J Am Chem Soc 138, 7496–7499, doi:10.1021/jacs.6b03836 (2016). [PubMed: 27280683]
12. Popp MW, Antos JM, Grotenbreg GM, Spooner E & Ploegh HL Sortagging: a versatile method for protein labeling. Nat Chem Biol 3, 707–708, doi:10.1038/nchembio.2007.31 (2007). [PubMed: 17891153]
13. Antos JM et al. A straight path to circular proteins. J Biol Chem 284, 16028–16036, doi:10.1074/jbc.M901752200 (2009). [PubMed: 19359246]
14. Antos JM et al. Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity. J Am Chem Soc 131, 10800–10801, doi:10.1021/ja902681k (2009). [PubMed: 19610623]
15. Freiburger L et al. Efficient segmental isotope labeling of multi-domain proteins using Sortase A. Journal of biomolecular NMR 63, 1–8 (2015). [PubMed: 26319988]
16. Bartels L, Ploegh HL, Spits H & Wagner K Preparation of bispecific antibody-protein adducts by site-specific chemo-enzymatic conjugation. Methods 154, 93–101, doi:10.1016/j.ymeth.2018.07.013 (2019). [PubMed: 30081077]

17. Harmand TJ et al. One-Pot Dual Labeling of IgG 1 and Preparation of C-to-C Fusion Proteins Through a Combination of Sortase A and Butelase 1. Bioconjug Chem 29, 3245–3249, doi:10.1021/acs.bioconjchem.8b00563 (2018). [PubMed: 30231608]
18. Piotukh K et al. Directed evolution of sortase A mutants with altered substrate selectivity profiles. J Am Chem Soc 133, 17536–17539 (2011). [PubMed: 21978125]
19. Chen L et al. Improved variants of SrtA for site-specific conjugation on antibodies and proteins with high efficiency. Scientific reports 6, 31899 (2016). [PubMed: 27534437]
20. Chen I, Dorr BM & Liu DR A general strategy for the evolution of bond-forming enzymes using yeast display. Proc Natl Acad Sci U S A 108, 11399–11404, doi:10.1073/pnas.1101046108 (2011). [PubMed: 21697512]
21. Dorr BM, Ham HO, An C, Chaikof EL & Liu DR Reprogramming the specificity of sortase enzymes. Proc Natl Acad Sci U S A 111, 13343–13348, doi:10.1073/pnas.1411179111 (2014). [PubMed: 25187567]
22. Selkoe DJ & Hardy J The amyloid hypothesis of Alzheimer's disease at 25 years. EMBO Mol Med 8, 595–608, doi:10.15252/emmm.201606210(2016).
23. Walsh D, Hartley D & Selkoe D The Many Faces of Aβ: Structures and Activity. Current Medicinal Chemistry - Immunology, Endocrine & Metabolic Agents 3, 277–291, doi:10.2174/1568013033483311 (2003).
24. Chen G. f. et al. Amyloid beta: structure, biology and structure-based therapeutic development. Acta Pharmacologica Sinica 38, 1205–1235, doi:10.1038/aps.2017.28 (2017). [PubMed: 28713158]
25. Brothers HM, Gosztyla ML & Robinson SR The Physiological Roles of Amyloid-β Peptide Hint at New Ways to Treat Alzheimer's Disease. Front Aging Neurosci 10, 118–118, doi:10.3389/fnagi.2018.00118 (2018). [PubMed: 29922148]
26. O'Brien RJ & Wong PC Amyloid Precursor Protein Processing and Alzheimer's Disease. Annual Review of Neuroscience, Vol 34 34, 185–204, doi:10.1146/annurev-neuro-061010-113613 (2011).
27. Roche J, Shen Y, Lee JH, Ying J & Bax A Monomeric Aβ1–40 and Aβ1–42 Peptides in Solution Adopt Very Similar Ramachandran Map Distributions That Closely Resemble Random Coil. Biochemistry 55, 762–775, doi:10.1021/acs.biochem.5b01259 (2016). [PubMed: 26780756]
28. O'Malley TT et al. Aβ dimers differ from monomers in structural propensity, aggregation paths and population of synaptotoxic assemblies. The Biochemical journal 461, 413–426, doi:10.1042/bj20140219 (2014). [PubMed: 24785004]
29. Boder ET & Wittrup KD Yeast surface display for screening combinatorial polypeptide libraries. Nature Biotechnology 15, 553–557, doi:10.1038/nbt0697-553 (1997).
30. Feldhaus MJ et al. Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. Nature Biotechnology 21, 163–170, doi:10.1038/nbt785 (2003).
31. Angelini A et al. in Yeast Surface Display: Methods, Protocols, and Applications Vol. 1319 Methods in Molecular Biology (ed Liu B) 3–36 (2015). [PubMed: 26060067]
32. Marraffini LA, Ton-That H, Zong Y, Narayana SVL & Schneewind O Anchoring of Surface Proteins to the Cell Wall of Staphylococcus aureus: A CONSERVED ARGININE RESIDUE IS REQUIRED FOR EFFICIENT CATALYSIS OF SORTASE A. Journal of Biological Chemistry 279, 37763–37770, doi:10.1074/jbc.M405282200 (2004).
33. Frankel BA, Tong Y, Bentley ML, Fitzgerald MC & McCafferty DG Mutational Analysis of Active Site Residues in the Staphylococcus aureus Transpeptidase SrtA. Biochemistry 46, 7269–7278, doi:10.1021/bi700448e (2007). [PubMed: 17518446]
34. Zong Y, Bice TW, Ton-That H, Schneewind O & Narayana SV Crystal structures of Staphylococcus aureus sortase A and its substrate complex. J Biol Chem 279, 31383–31389, doi:10.1074/jbc.M401374200 (2004). [PubMed: 15117963]
35. Suree N et al. The structure of the Staphylococcus aureus sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. J Biol Chem 284, 24465–24477, doi:10.1074/jbc.M109.022624 (2009). [PubMed: 19592495]
36. Kruger RG, Dostal P & McCafferty DG Development of a high-performance liquid chromatography assay and revision of kinetic parameters for the Staphylococcus aureus sortase

transpeptidase SrtA. Anal Biochem 326, 42–48, doi:10.1016/j.ab.2003.10.023 (2004). [PubMed: 14769334]

37. Ton-That H, Mazmanian SK, Faull KF & Schneewind O Anchoring of surface proteins to the cell wall of Staphylococcus aureus. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. J Biol Chem 275, 9876–9881, doi:10.1074/jbc.275.13.9876 (2000). [PubMed: 10734144]
38. Pawlowski M, Meuth SG & Duning T Cerebrospinal Fluid Biomarkers in Alzheimer's Disease-From Brain Starch to Bench and Bedside. Diagnostics (Basel) 7, 42, doi:10.3390/diagnostics7030042 (2017).
39. Baranello RJ et al. Amyloid-beta protein clearance and degradation (ABCD) pathways and their role in Alzheimer's disease. Curr Alzheimer Res 12, 32–46, doi:10.2174/1567205012666141218140953 (2015). [PubMed: 25523424]
40. Hong W et al. Diffusible, highly bioactive oligomers represent a critical minority of soluble Aβ in Alzheimer's disease brain. Acta Neuropathol 136, 19–40, doi:10.1007/s00401-018-1846-7 (2018). [PubMed: 29687257]
41. Mengel D et al. Dynamics of plasma biomarkers in Down syndrome: the relative levels of Aβ42 decrease with age, whereas NT1 tau and NfL increase. Alzheimer's Research & Therapy 12, 27, doi:10.1186/s13195-020-00593-7 (2020).
42. Isaacs AM, Senn DB, Yuan M, Shine JP & Yankner BA Acceleration of amyloid beta-peptide aggregation by physiological concentrations of calcium. The Journal of biological chemistry 281, 27916–27923, doi:10.1074/jbc.M602061200 (2006). [PubMed: 16870617]
43. Colvin MT et al. Atomic Resolution Structure of Monomorphic Abeta42 Amyloid Fibrils. J Am Chem Soc 138, 9663–9674, doi:10.1021/jacs.6b05129 (2016). [PubMed: 27355699]
44. Walsh DM et al. A facile method for expression and purification of the Alzheimer's disease-associated amyloid beta-peptide. FEBS J 276, 1266–1281, doi:10.1111/j.1742-4658.2008.06862.x (2009). [PubMed: 19175671]
45. O'Malley TT, Witbold WM 3rd, Linse S & Walsh DM The Aggregation Paths and Products of Aβ42 Dimers Are Distinct from Those of the Aβ42 Monomer. Biochemistry 55, 6150–6161, doi:10.1021/acs.biochem.6b00453 (2016). [PubMed: 27750419]
46. Bacon K, Burroughs M, Blain A, Menegatti S & Rao BM Screening Yeast Display Libraries against Magnetized Yeast Cell Targets Enables Efficient Isolation of Membrane Protein Binders. ACS Combinatorial Science 21, 817–832, doi:10.1021/acscombsci.9b00147 (2019). [PubMed: 31693340]
47. Nguyen GK et al. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. Nat Chem Biol 10, 732–738, doi:10.1038/nchembio.1586 (2014). [PubMed: 25038786]
48. Yang R et al. Engineering a catalytically efficient recombinant protein ligase. J Am Chem Soc 139, 5351–5358 (2017). [PubMed: 28199119]
49. Nikghalb KD et al. Expanding the Scope of Sortase-Mediated Ligations by Using Sortase Homologues. Chembiochem 19, 185–195 (2018). [PubMed: 29124839]
50. Tang S, Xuan B, Ye X, Huang Z & Qian Z A modular vaccine development platform based on sortase-mediated site-specific tagging of antigens onto virus-like particles. Scientific reports 6, 25741 (2016). [PubMed: 27170066]

Methods-only References

51. Meyer AJ, Ellefson JW & Ellington AD Library generation by gene shuffling. Curr Protoc Mol Biol 105, Unit 15 12, doi:10.1002/0471142727.mb1512s105 (2014). [PubMed: 24510437]
52. Chao G et al. Isolating and engineering human antibodies using yeast surface display. Nature Protocols 1, 755–768, doi:10.1038/nprot.2006.94 (2006). [PubMed: 17406305]
53. Killenberg PG & Dukes DF Coenzyme A derivatives of bile acids-chemical synthesis, purification, and utilization in enzymic preparation of taurine conjugates. Journal of lipid research 17, 451–455 (1976). [PubMed: 9462]
54. Mofid MR, Finking R, Essen LO & Marahiel MA Structure-Based Mutational Analysis of the 4'-Phosphopantetheinyl Transferases Sfp from Bacillus subtilis: Carrier Protein Recognition and

Reaction Mechanism. Biochemistry 43, 4128–4136, doi:10.1021/bi036013h (2004). [PubMed: 15065855]

55. Tropea JE, Cherry S & Waugh DS Expression and purification of soluble His(6)-tagged TEV protease. Methods Mol Biol 498, 297–307, doi:10.1007/978-1-59745-196-3_19 (2009). [PubMed: 18988033]
56. Gietz RD & Schiestl RH High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protocols 2, 31–34 (2007). [PubMed: 17401334]
57. Walsh DM, Lomakin A, Benedek GB, Condron MM & Teplow DB Amyloid β-Protein Fibrillogenesis: DETECTION OF A PROTOFIBRILLAR INTERMEDIATE. Journal of Biological Chemistry 272, 22364–22372, doi:10.1074/jbc.272.35.22364 (1997).

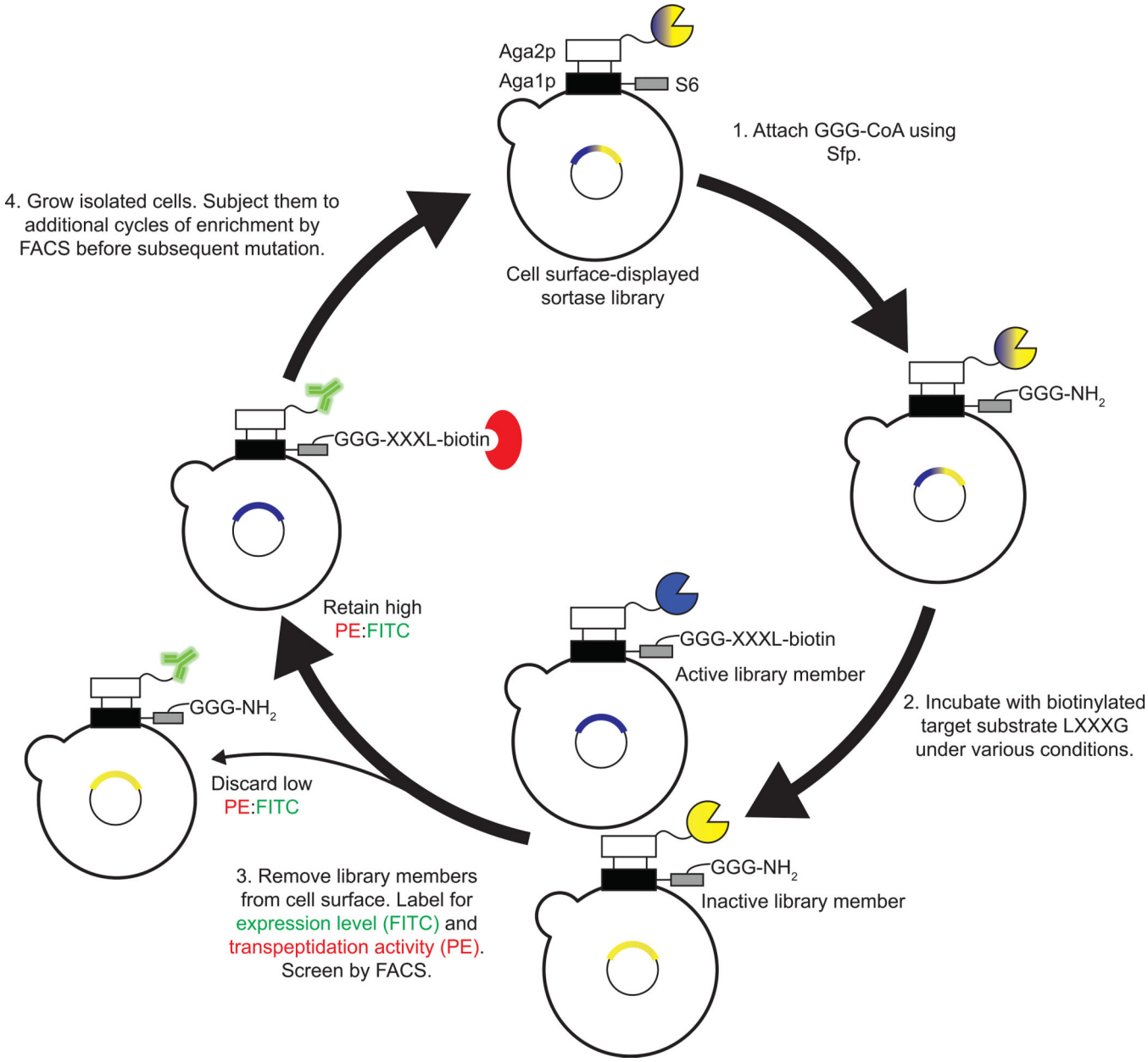


Figure 1. Yeast display strategy for sortase evolution.

A population of yeast displays a library of ~10^7 SrtA variants. 1) Triglycine is conjugated to the surface of each cell with Sfp phosphopathetheinyl transferase. 2) The cells are incubated with biotinylated target substrate and non-biotinylated off-target substrates. 3) After allowing the SrtA variants to catalyze transpeptidation between triglycine and the added substrates, cells are washed and the SrtA variants are removed from their surfaces using TEV protease. Cells are labelled with an anti-HA antibody (in green) to quantify sortase expression and streptavidin-PE (in red) to quantify transpeptidation between triglycine and the positive selection substrate. Active sortase variants (blue) have higher on-target transpeptidation per unit expression than inactive variants (yellow) or promiscuous variants

and can be isolated by fluorescence activated cell sorting (FACS). 4) Collected cells were grown, reinduced, and further enriched for target recognition.

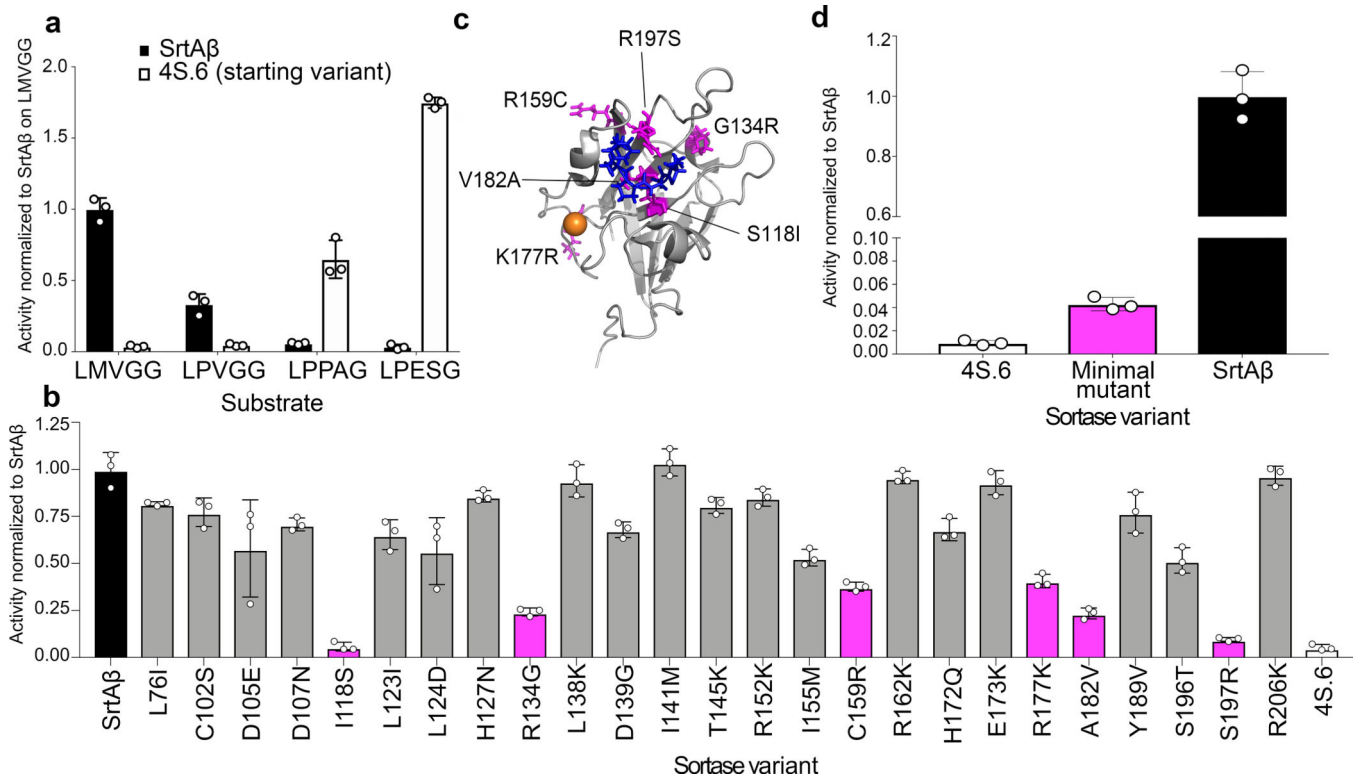


Figure 2. Activity profile and mutational analysis of SrtAβ.

(**a**) The evolved SrtAβ and the starting enzyme 4S.6 were displayed on yeast and assayed for their ability to catalyze transpeptidation on different substrates. (**b**) SrtAβ, 4S.6, and all 25 single-reversion mutants were displayed on yeast and assayed for their ability to catalyze transpeptidation between triglycine and Btn-LMVGG. Reversion mutants with activity less than half that of SrtAβ are highlighted in pink. (**c**) The predicted locations of the six reversion mutations that reduce SrtAβ activity by >50% are shown in pink on the NMR solution structure of wild-type *S. aureus* SrtA (PDB: 2KID). An LPXTG substrate analogue is blue, and the calcium ion required for activity is orange. Residues 118, 182, and 197 are part of the substrate binding pocket, while other residues are further from the active site. (**d**) The activity of 4S.6, a minimal mutant (4S.6 with S118I, G134R, R159C, K177R, V182A, and R197S mutations), and SrtAβ on Btn-LMVGG were compared by flow cytometry. Addition of these six mutations to 4S.6 improves activity on the LMVGG substrate, but is insufficient to confer the level of activity displayed by SrtAβ, highlighting the importance of other mutations. All graphs represent the mean of three replicates ± standard deviation. Activity is defined as the ratio of cell surface biotinylation (PE) to sortase expression level (FITC).

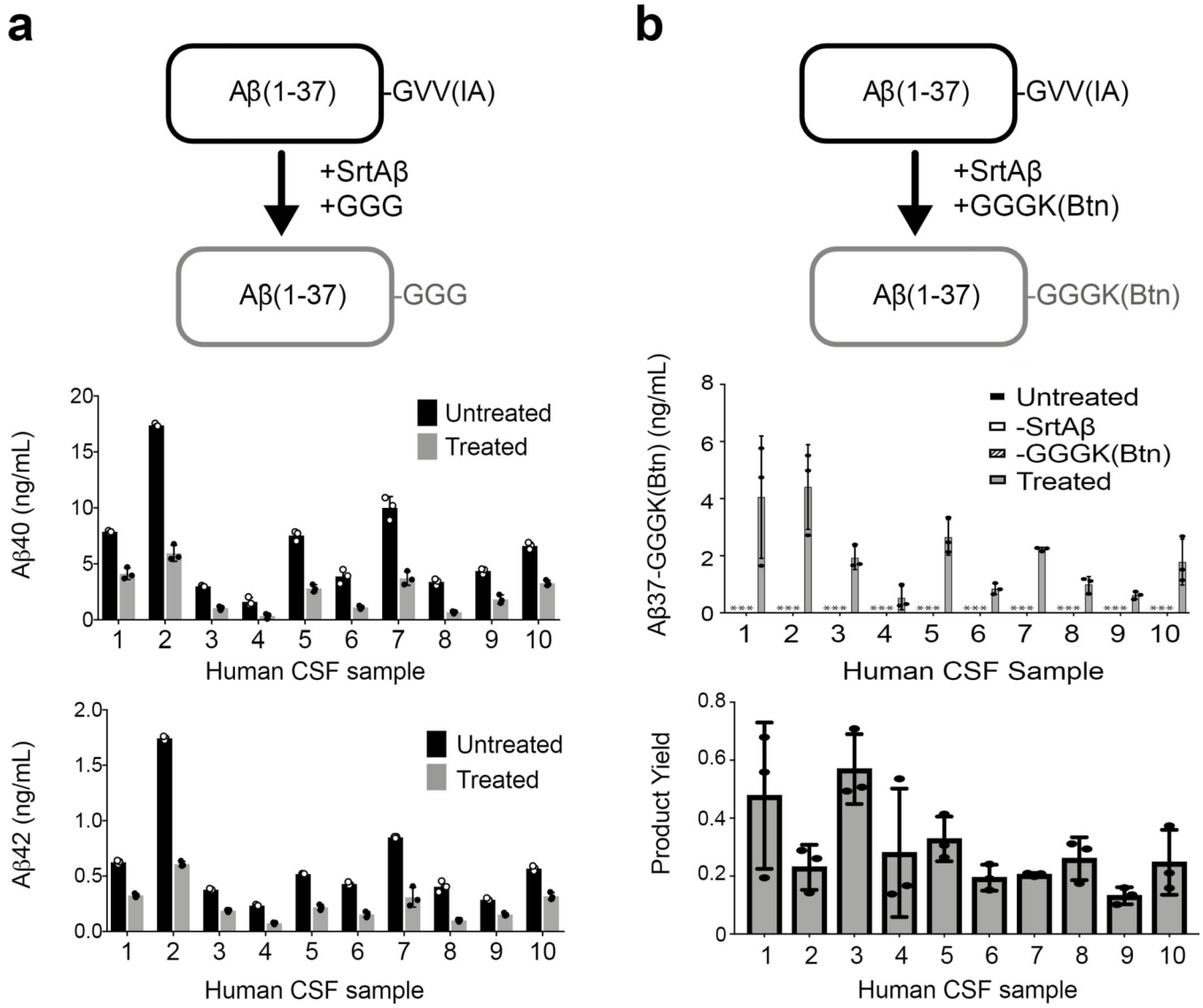


Figure 3. Sortase labeling of endogenous Aβ in human cerebrospinal fluid.

(**a**) Transpeptidation of Aβ40 or Aβ42 with GGG should yield Aβ37-GGG, which is not detected by Aβ40- and Aβ42-specific ELISAs. Treatment of CSF specimens with SrtAβ and GGG caused a significant reduction in ELISA-measured levels of both Aβ40 and Aβ42. (**b**) Transpeptidation of Aβ40 or Aβ42 with GGGK(Btn) yields Aβ37-GGGK(Btn), which can be detected through its affinity handle. Detectable levels of transpeptidation product are observed in all 10 CSF samples. Importantly, no product was observed (* indicates below limit of detection) in the absence of SrtAβ or GGGK(Btn). Product yield is defined as the amount of product detected divided by the amount of Aβ40 + Aβ42 measured in each sample. For each labeling experiment, all reactions were set up in triplicate. Bars represent the mean of three replicates ± standard deviation. The GGG labeling experiment was performed once. The GGGK(Btn) labeling experiment was performed twice. The data presented are representative of both attempts.

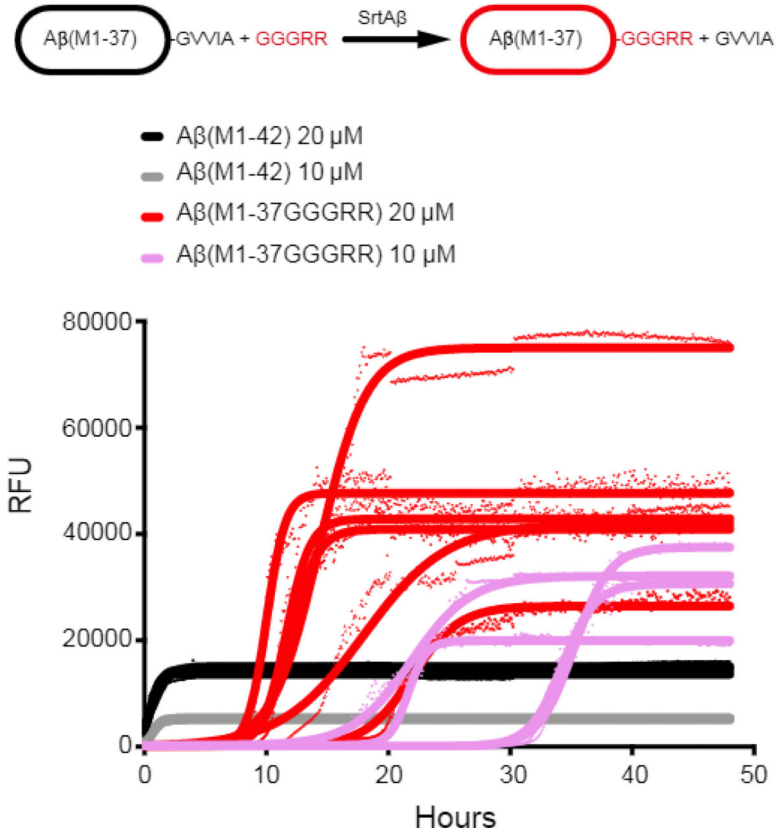


Figure 4. Aggregation inhibition of Aβ42 with SrtAβ.

Thioflavin T binding was used to monitor the aggregation of AβM1–42 and AβM1–37GGGRR. Data points from the time-course are shown for each replicate (n = 3 for AβM1–42, n = 4 for 10 μM Aβ(M1–37GGGRR), and n = 6 for 20 μM Aβ(M1–37GGGRR)) and curves fitted to each replicate by Boltzmann equation are indicated. The initiation of aggregation of Aβ(M1–37GGGRR) monomer was greatly retarded compared to AβM1–42, with an average $t_{1/2}$ = 14.6 hours at 20 μM and 28.3 hours at 10 μM (compared to 0.6 hours and 0.7 hours for AβM1–42 at 20 μM and 10 μM, respectively).

Table 1.
Mutations observed in evolved sortase variants.

Mutations in representative clones from rounds 1–16 are shown relative to the evolutionary starting sequence, 4S.6. These clones were the most abundant sequences in the evolving pool at the end of their rounds, with the exceptions of clone 8.5-H3 and SrtAβ, which were identified by single clone FACS at the end of round 8 and round 16, respectively.

4S.6	R1	R2	R4	R5	R7	R8	8.5-H3	R9	R10	R11	R12	R13	SrtAβ
K62						R	R	R	R	R	R	R	
A73									V		V		
I76						L		L	L	L	L	L	L
R94	P	Y	Y	Y	Y								
S102													C
E105												D	D
N107							D	D	D	D	D	D	D
S118	I	I	I	I	I	I	I	I	I	I	I	I	I
A122		W											
I123				L	L	L	L	L	L	L	L	L	L
D124		G	L	L	L	L	L	L	L	L	L	L	L
N127						Y			Y	H	H	H	H
G134	R	R	R	R	R	R	R	R	R	R	R	R	R
K138			I	I	I	L	I	I	L	L	L	L	L
G139													D
M141													I
K145						T			T	T	T	T	T
G147												C	
N148											S		
K152													R
M155										I	I	I	I
S157								R					
R159						C		H	C	C	C	C	C
K162													R
D170										E	E	E	
Q172									H	H	H	H	H
K173					E		E	E	E	E	E	E	E
K177							R	R					R
V182			A	A	A	A	A	A	A	A	A	A	A
V189	F	F	F	F	F	F	I	I	F	F	F	F	Y
T196			S	S	S	S	S	S	S	S	S	S	S
R197			S	S	S	S	S	S	S	S	S	S	S
K206									E	E			R