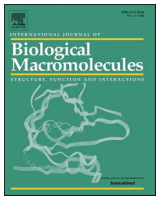
International Journal of Biological Macromolecules 272 (2024) 132946

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac



Utilization of low-stability variants in protein evolutionary engineering

Mitsutoshi Wakisaka, Shun-ichi Tanaka, Kazufumi Takano *

Department of Biomolecular Chemistry, Kyoto Prefectural University, Sakyo-ku, Kyoto 606-8522, Japan

ARTICLE INFO

Keywords:
Evolutionary engineering
Protein stability
Protein activity

ABSTRACT

Evolutionary engineering involves repeated mutations and screening and is widely used to modify protein functions. However, it is important to diversify evolutionary pathways to eliminate the bias and limitations of the variants by using traditionally unselected variants. In this study, we focused on low-stability variants that are commonly excluded from evolutionary processes and tested a method that included an additional restabilization step. The esterase from the thermophilic bacterium *Alicyclobacillus acidocaldarius* was used as a model protein, and its activity at its optimum temperature of 65 °C was improved by evolutionary experiments using random mutations by error-prone PCR. After restabilization using low-stability variants with low-temperature (37 °C) activity, several re-stabilizing variants were obtained from a large number of variant libraries. Some of the restabilized variants achieved by removing the destabilizing mutations showed higher activity than that of the wild-type protein. This implies that low-stability variants with low-temperature activity can be re-evolved for future use. This method will enable further diversification of evolutionary pathways.

1. Introduction

Protein engineering strategies include rational and semi-rational approaches and directed evolution methods [1,2]. Directed evolution is a widely used technique that does not require structural information and can lead to the discovery of unexpected and effective mutations [3–7]. By modifying enzyme properties, directed evolution experiments using random mutations sometimes can detect mutations that are distant from the active site [8].

The directed evolution of enzymes is a method of modifying enzymes to suit their desired functions in a manner that mimics natural selection [9,10]. The process consists of (i) diversification of the target gene sequence, (ii) screening of the expressed variant proteins using a specific index, and (iii) repeatedly using the sequence that encodes the selected variant as a new template for the next cycle to obtain better variants. However, the continuous evolution of an enzyme property could reduce the improvement in functionality over time [11]. This is because selecting variants using a single target index (selective pressure) involves only a small range of the large variant sequence space. This is shown as route A in the fitness landscape in Fig. 1a. Furthermore, amino acid substitutions are often destabilizing [12–14]. Thus, as mutations accumulate, variants that cannot form structures are removed from the evolutionary pathway [15,16]. To overcome these issues, a method was devised to introduce a stability index, i.e., adopting stability as a selective pressure, or screening based on stability [17–21], pool stable variants regardless of the objective index, and use them as scaffolds for adding mutations. This strategy has allowed for the evaluation of the effects of a wide range of variants and the discovery of highly functional variants, as shown in routes B and C in Fig. 1a [22]. Therefore, we hypothesized that the evolutionary pathway could be further diversified with variants by adding a step to re-stabilize the low-stability variants that are conventionally eliminated.

In this study, we investigated the evolution of proteins to improve their activity through mutation. An esterase from the thermophilic bacterium *Alicyclobacillus acidocaldarius* (Aac-Est) was used as the model protein [23,24]. Aac-Est is a monomer consisting of 310 amino acid residues that has a molecular weight of 34,800, an optimum temperature of approximately 65 °C, and a T_m value of 91 °C [25–29]. Previous studies have conducted activity screens using the Aac-Est wild-type protein as a template (approximately one to three amino acid substitutions), although no variants with enhanced activity were found at 65 °C [20,24]. This means that the Aac-Est wild-type is located at the local peak in the fitness landscape shown in Fig. 1b. Furthermore, when these variants that decreased activity were screened again using the stability-maintaining variants as templates, activity-enhancing variants were obtained, but not from the variants with decreased stability [20]. This indicates that a minimum threshold of stability is required to go "downhill" followed by "uphill" in sequence space. Therefore, we

* Corresponding author.
E-mail address: takano@kpu.ac.jp (K. Takano).

https://doi.org/10.1016/j.ijbiomac.2024.132946
Received 3 May 2024; Received in revised form 3 June 2024; Accepted 4 June 2024
Available online 5 June 2024

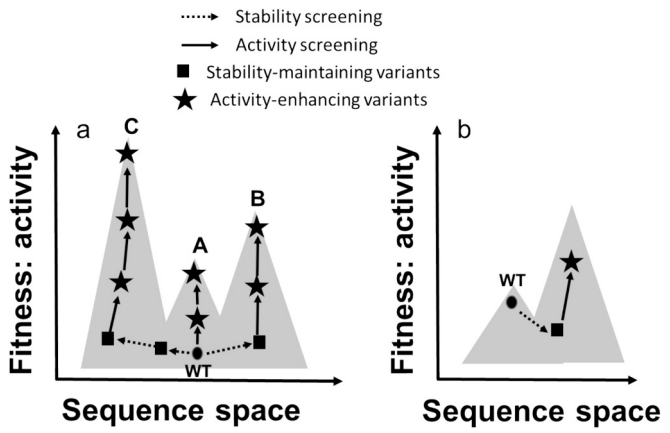


Fig. 1. Fitness landscapes of the activity against the sequence space. (a) Wild-type protein (circle) is located near the foot of a low mountain. Route A: the result of conventional evolution experiments. Routes B and C: the results of evolutionary experiments adding stability screening. (b) Wild-type protein (circle) is located near the top of a low mountain. Stability-maintaining variants (squares) are obtained by stability screenings (dashed arrows). Activity-enhancing variants (stars) are obtained by activity-screenings (linear arrows).

attempted to obtain variants with improved activity by introducing a stability index and verifying the feasibility of a method that added a restabilization step for low-stability variants that did not meet the selection criteria (variants below the minimum threshold of stability).

2. Materials and methods

2.1. Random and site-specific mutagenesis

A plasmid for Aac-Est overexpression, designated pET28a/Aac-Est, was constructed as previously reported [20]. In this study, plasmid pET28a/AE_WT, which was deleted by 60 bp upstream of the Aac-Est gene from pET28a/Aac-Est, was used as a template. With this design, the primers covered the Shine-Dalgarno sequence involved in translation upstream of the target, and only the Aac-Est gene was targeted for mutagenesis.

Random mutations were introduced by error-prone PCR using rTaq polymerase [30–32]. The reaction mixture contained 0.3 mM $MnCl_2$, optimized to generate an error frequency of one to four substitutions per gene. When there were multiple templates, they were mixed after subjecting each plasmid to error-prone PCR. Site-specific mutations were performed by inverse PCR using Q5® High-Fidelity DNA Polymerase with destabilized mutant plasmids as templates. Mutated genes amplified by error-prone PCR were digested using *Xba*I and *Sac*I-HF and ligated to the corresponding sites of pET28a at 16 °C for overnight. The mutant plasmids were transformed into *Escherichia coli* JM109(DE3)-competent cells. Transformed *E. coli* JM109(DE3) cells were grown on LB plates with 50 μg/mL kanamycin, 1 % tributyrin and 0.1 % Tween 20 (tributyrin plate) at 37 °C for overnight.

2.2. Activity and stability evaluation by halo assay

Among the single colonies obtained after overnight plate incubation, those that had degraded tributyrin and formed a halo were selected and inoculated onto LB plates containing 2 % gellan gum (gellan gum plate). As a comparison, colonies obtained by similar transformation using pET28a and pET28a/AE_WT were also inoculated. Inoculation was done on two or three gellan gum plates to evaluate halo at different temperatures for the same colonies. LB agar plates for bacteria preservation (replica plate) were also inoculated in the same way and these plates were incubated at 37 °C for overnight.

The next day, gellan gum plates were layered with 1 % tributyrin, 1 % Tween 20, 2 % gellan gum and incubated at 65 °C, the optimal temperature for Aac-Est wild-type protein, for 3 h [33]. One set of plates was heat treated at 80 °C for 30 min before layering (heat treatment, 80 °C → 65 °C). Colonies that formed a halo on the 80 °C → 65 °C plate were determined to express the stability maintaining variant (or Aac-Est wild-type protein). The plasmids of the corresponding colonies were extracted from the colonies in the replica plates. Sequence analysis was contracted to Eurofin Genomics, Inc. The variants in which amino acid substitutions had been introduced were selected and further random mutations were introduced using the corresponding plasmids as templates. The above random mutations and halo formation were carried out until the 5th round. Colonies that formed a halo at 37 °C but not sufficiently at 65 °C were judged to express low-stability variants. Plasmids were extracted and used for re-stabilization experiments from the low-stability variants.

In the acquisition of the re-stabilizing variants, a random mutation was introduced into the plasmid encoding the low-stability variants. *E. coli* JM109(DE3) was transformed with the random mutant plasmids. Transformed *E. coli* JM109(DE3) cells were grown on LB plates with 50 μg/mL kanamycin at 37 °C for overnight. After overnight incubation, the plates were directly overlaid with 1 % tributyrin, 1 % Tween 20, and 2 % gellan gum and incubated at 65 °C for 3 h. The colonies that formed halo were judged to express the variant that had been re-stabilized to an active level at the optimal temperature.

Since the colonies were not inoculated onto gellan gum plates or replica plates, many colonies from the post-transformation LB plates could be directly used for halo assay evaluation. This method is effective because the number of re-stabilized variants is very limited. However, since the bacteria were killed by direct layering and incubation at 65 °C, the portion with halo-forming colonies was pinched off from the layered plate, mixed with 20 μL of MilliQ water, centrifuged to precipitate solids, and the supernatant was used for transformation of *E. coli* strain JM109(DE3). Each colony after plate culture was inoculated and halo evaluated again.

2.3. Protein expression and cell-free extract

The relevant *E. coli* were cultured at 37 °C in LB medium containing 50 μg/mL kanamycin. Expression was induced with 0.25 mM isopropyl-β-D-thiogalactopyranoside (IPTG) when cells reached an OD_{600} of 0.5. After overnight growth at 18 °C, the medium was aliquoted to match the amounts of bacteria by the OD_{600} measurement, and cells were collected by centrifugation and resuspended in 50 mM Tris-HCl (pH 8.0) containing 10 mM EDTA and 150 mM NaCl. To prepare cell-free extracts of proteins, cells were lysed with 1 mg/mL lysozyme for 30 min at 37 °C and 0.1 % TritonX-100 for 1 h on ice. After centrifugation for 45 min, the supernatant was collected and used as a cell-free extract of enzyme. The expression level of each protein was analyzed using 15 % SDS-PAGE with ImageJ standardization at several concentrations of BSA.

2.4. Activity assay

Activity after the cleavage of the substrate *p*-nitrophenyl caproate (pNC) was measured for 3 min at 65 °C with a cell-free extract of enzyme [20]. The reaction was stopped by adding 10 % SDS solution. The reaction mixture contained 0.2 mM pNC, 40 mM phosphate buffer (pH 7.1), 0.09 % (*v*/v) gum arabic, 0.36 % (v/v) Triton X-100, and 2 % (v/v) 2-propanol. The amount of product was measured using a spectrophotometer at 405 nm. Activity was normalized by the expression level.

3. Results and discussion

In this study, three evolutionary engineering approaches, as shown in Fig. 2, involving error-prone PCR were used to obtain Aac-Est variants with enhanced activity at 65 °C. We first used the classical evolutionary engineering method (activity screening) of selecting highly active

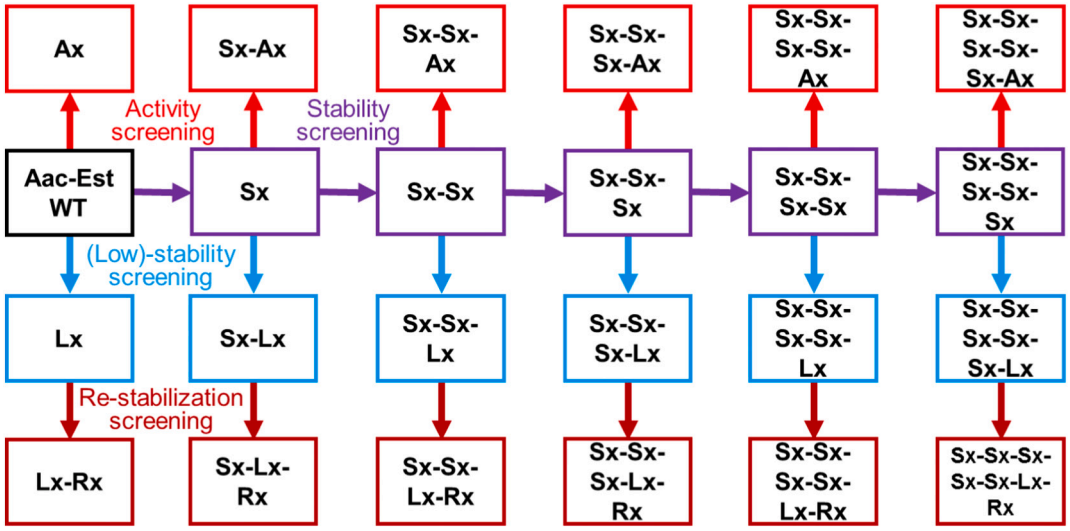


Fig. 2. Screening steps in this work. Variants obtained in the activity, stability, (low-)stability and re-stabilization screenings are denoted by Ax, Sx, Lx and Rx, and those obtained in the activity screening after one round of stability screening are denoted as Sx-Ax.

variants from a randomly generated group by activity-based screening using the Aac-Est wild-type protein as a template. We then used stability-based screening (stability and activity screening) to select stability-maintaining variants from random mutations using the wild-type Aac-Est protein as a template. These variants exhibited halo activity at 65 °C after heat treatment at 80 °C. The resulting stability-maintaining variants were used as templates for introducing additional random mutations to obtain further stability-maintaining variants. This process was performed for five rounds. We then introduced random mutations into the stability-maintaining variants obtained in rounds 1 to 5 and selected highly active variants by activity-based screening [22]. Finally, the random mutations obtained from Aac-Est wild-type protein or stability-maintaining variants were used to select low-stability variants that lost their activity at 65 °C, but showed activity at 37 °C [(low-)stability and restabilization screening]. Random mutations were introduced into these low-stability variants to obtain active restabilized variants at 65 °C. Variants obtained in the activity, stability, (low-)stability, and restabilization screenings were denoted as Ax, Sx, Lx, and Rx, respectively, and those obtained in the activity screening after one round of stability screening were denoted as Sx-Ax.

In the halo evaluation in this study, halo formation was confirmed on gellan gum plates without inducer, such as IPTG. We speculate that this is because the degradation reaction by Aac-Est expressed at the basal level was sufficient to form halos even without induction of protein expression by inducer.

3.1. *Activity screening from random mutations of Aac-Est wild-type protein*

Previous studies using the Aac-Est wild-type protein as a template found no variant with increased activity at 65 °C in the first round of screening [20,24]. This indicates that the wild-type protein at 65 °C is located near the top of a specific mountain (localized peak), where it is difficult to obtain a highly active variant by activity screening. The Aac-Est wild-type protein is stabilized by ionic bonds on the surface of the molecule and destabilized by the introduction of mutations on the surface and inside the molecule [24]. Consequently, the structure could not be maintained at 65 °C and the activity decreased. However, these variants have shown activity at 40 °C in various cases [20,24].

In this study, we attempted to obtain highly active variants of the wild-type protein using an activity screening method described in previous studies [20,24] with the introduction of a halo activity check at 65 °C. The previous method involved the random selection of variants and subsequent measurement of their activity. The results included many variants with a complete loss of activity at 65 °C. The halo activity check prevents selection of variants with diminished activity.

The relative activity of the variants to the wild-type protein at 65 °C ranged from 0.54 to 1.2, as shown in Fig. 3, in which each variant is labeled from A1 to Ax. The variants are arranged in descending order of their relative activities. Although several variants showed activities higher than those of the wild-type protein, when the error of ±0.1 in activity for the same variant was considered, the only highly active variant was A1 with a relative activity of 1.2. This result is similar to those of previous studies, which returned no highly active variants [20,24]. The introduction of the halo activity check may have allowed us to identify the slightly more active variant, A1, which had the K296E substitution (Table 1). The purpose of this study was to verify the use of low-stability variants in protein evolutionary engineering; therefore, we did not discuss the mechanisms by which individual amino acid substitutions affect activity and stability. To properly examine the effects of multiple amino acid substitutions introduced by random mutations, it is necessary to create single variants and analyze their structures.

The proportion of measured samples with activity-enhancing variants was defined as evolvability [16,20], and this evolvability was 6.7 % when the Aac-Est wild-type protein was used as a template. This low result suggests that it is difficult to obtain highly active variants from wild-type proteins using activity screening alone and other indexes are

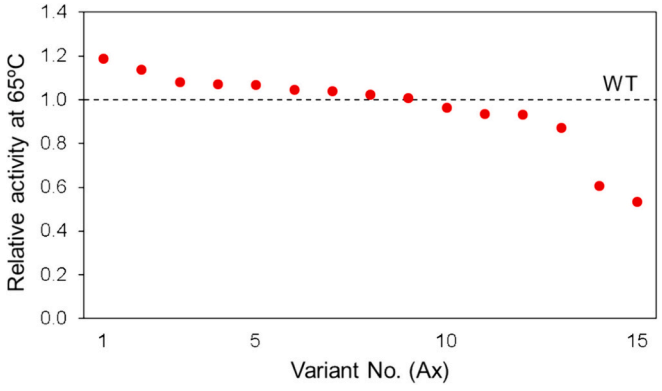


Fig. 3. Relative activity of variants obtained from Aac-Est wild-type protein by activity screening. Each variant is labeled to Ax. The variants are arranged in descending order of relative activity.

Table 1
Character of activity-enhancing variant obtained by activity screening.

Variant	Substitution	Relative activity to WT at 65 °C
A1	K296E	1.2

required. Evolvability is further discussed in Section 3.2 in conjunction with the results of the stability and activity screenings.

3.2. Stability and activity screenings

In our study, it was difficult to obtain Aac-Est variants with enhanced activity using activity screening from random mutations and the wild-type protein as a template. However, an improved variant (relative activity of 1.5 to the wild-type protein) was identified by activity screening from a variant with decreased activity (relative activity of 0.6) while maintaining stability [20]. Therefore, it is likely that the fitness landscape involved descending the mountain, moving to the base of the mountain at a different peak, and then re-ascending the mountain. Thus, stability screening can reduce the bias and restriction of variants to be explored, thereby allowing for new useful variants to be obtained and a diversified evolutionary pathway to be created. This has been demonstrated by the acquisition of highly active esterase variants from the hyperthermophilic archaeon *Sulforobus tokodaii* (Sto-Est) [22]. Five rounds of stability screening were successfully conducted in that study to identify the more potent variants. Therefore, we used activity screening from multiple stability screenings to obtain activity-enhancing Aac-Est variants.

First, from the random mutations created using the wild-type protein as a template, we obtained six stability-maintaining variants that had halo activities at 65 °C after heat treatment at 80 °C. Subsequently, 11 stability-maintaining variants were acquired in the same manner, using each of these stability-maintaining variants as a template. This operation was continued for up to five rounds, resulting in 44 stability-maintaining variants. The amino acid substitutions of these variants per round ranged from one to four, with an average of 1.5, and are listed in Table 2. Up to ten substitutions and five rounds of random mutations were performed with three variants. The amino acid substitution T164A overlapped with the stability-maintaining variants obtained in a previous study [20]. This indicates that the same amino acid substitution was reproducibly selected as the stability-maintaining variant during the stability screening. Various stability-maintaining variants were also identified.

The relative activity of the stability-maintaining variants S1–6 relative to the wild-type protein at 65 °C ranged from 0.82 to 1.1. This is likely because the halo activity at 65 °C was assessed, and variants with extremely reduced activity were not selected. Some of the acquired stability-maintaining variants showed greater activity than the wild-type protein (see below).

Using activity screening, we then attempted to obtain highly active variants from random mutations of the stability-maintaining variants derived in each round. The relative activity of the 15 stability-maintaining variants obtained in the first round of stability screening to the wild-type protein at 65 °C ranged from 0.51 to 1.4, as shown in Fig. 4a. Of these variants, four showed greater activities (relative activity of $\geq$1.2) than the wild-type protein; their amino acid substitutions are listed in Table 3. The results of stability-screening rounds 2–5 are shown in Figs. 4b-e and Table 3. Twenty activity-enhancing variants were selected.

Of these, the S4-S4, S4-S2-S1, and S4-S4-S1-S6 variants were considered stability-maintaining variants from the stability screening. These variants were likely derived from stability-maintaining variants from the preceding round. For example, S4-S4 could be the result of adding four additional amino acid substitutions to the stability-maintaining variant S4. Since S4-S4 is a variant from the second round of stability screening, it was treated in this study as a group of variants from the second round.

It is difficult to obtain a highly active variant by random mutation of the Aac-Est wild-type protein. However, through stability screening, we were able to successfully identify 20 activity-enhancing variants. Furthermore, a variant (S4-S2-S1-A1) was obtained that had higher activity (relative activity of 1.7) than that of the high activity variant (relative activity of 1.5) reported in a previous study [20]. This indicates that cumulative stability screening was effective and consistent with the

Table 2
Characters of stability-maintaining variants obtained by stability screenings.

Variants	Substitutions
S1	Y129F
S2	T226S
S3	V40A
S4	D52E
S5	L225P
S6	L25H
S2-S1	A124T T226S
S2-S2	K42R K120R Q217H T226S I281V
S2-S3	F51L T226S P253S
S3-S1	V40A R66S E199D L222Q
S3-S2	V40A W213R V233I
S4-S1	V45E D52E
S4-S2	K42R D52E
S4-S3	K23E S32P D52E I274T
S4-S4	D52E A140E L222P S239R A249V
S4-S5	Q8R L14H D52E
S5-S1	D103V N220D L225P
S2-S1-S1	V40A A124T Q133R T226S
S2-S3-S1	F51L R138H T226S P253S
S2-S3-S2	F51L T164A T226S P253S
S4-S2-S1	K42R D52E L90I T295M
S4-S4-S1	D52E A140E L222P S239R A249V K271E
S5-S1-S1	D103V N220D L225P D256G
S2-S3-S2-S1	F51L T164A S165C T226S P253S F277Y
S4-S2-S1-S1	K42R D52E L90I A106T A123V T295M
S4-S4-S1-S1	D52E L100F A140E L222P S239R A249V K271E
S4-S4-S1-S2	D52E A140E L222P P236L S239R A249V K271E
S4-S4-S1-S3	P39H D52E A140E S221G L222P S239R A249V K271E
S4-S4-S1-S4	D52E E91G A140E L222P S239R A249V K259E K271E
S4-S4-S1-S5	D52E A140E A147D L222P S239R A249V K271E D307G
S4-S4-S1-S6	F51L D52E A140E L222P S239R A249V K271E I274V
S5-S1-S1-S1	L36P K42R D103V N220D L225P D256G
S4-S2-S1-S1-S1	K42R D52E L90I A106T A123V N266Y T295M
S4-S2-S1-S1-S2	K42R D52E L90I A106T A123V L132F T295M
S4-S4-S1-S1-S1	R16S D52E L100F A140E L222P S239R A249V K271E
S4-S4-S1-S2-S1	N15D D52E A140E L222P P236L S239R A249V K271E
S4-S4-S1-S2-S2	D52E D103G A140E A201E L222P P236L S239R A249V K271E
S4-S4-S1-S3-S1	P39H D52E M64V A140E S221G L222P S239R A249V K271E N276Y
S4-S4-S1-S3-S2	P39H D52E A140E S221G L222P S239R A249V K267E K271E
S4-S4-S1-S4-S1	R16H P39H D52E E91G A140E L222P S239R A249V K259E K271E
S4-S4-S1-S5-S1	D52E A140E A147D T164A L222P S239R A249V K271E
S4-S4-S1-S6-S1	F51L D52E A140E L222P S239R A249V N266S K271E I274V
S4-S4-S1-S6-S2	Q8R F51L D52E P72A A140E L222P S239R A249V K271E I274V
S4-S4-S1-S6-S3	A19S F51L D52E A140E L222P S239R A249V K271E I274V

The underlined amino acid substitutions had already occurred in the template.

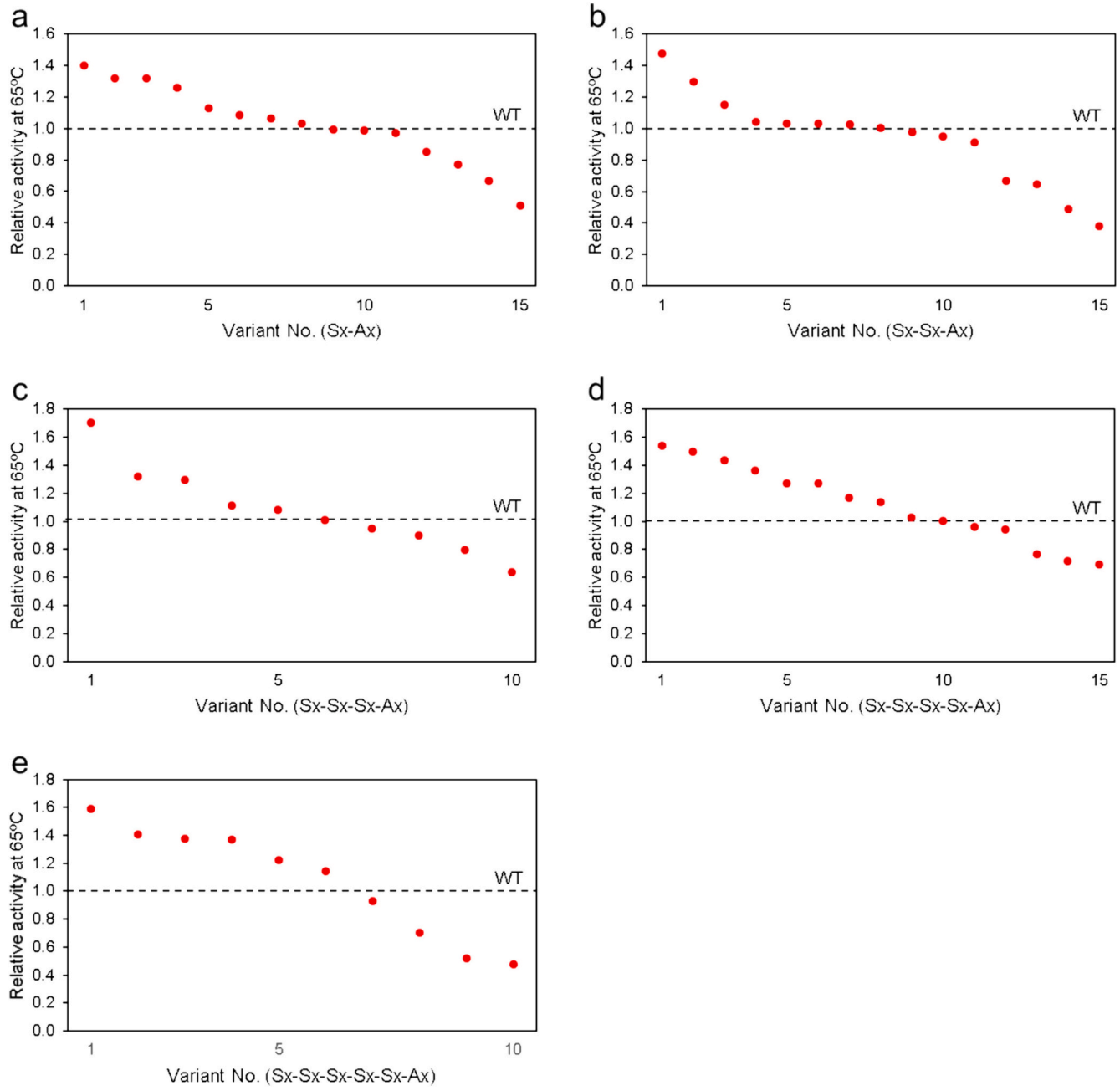


Fig. 4. Relative activity of variants obtained by activity screening after stability screenings. Results using the stability-maintaining variants from stability screening (a) first-, (b) second-, (c) third-, (d) forth- and (e) fifth-rounds. The figure style is the same as Fig. 2.

Sto-Est results.

The evolvability of each round is summarized in Table 4, which reveals higher values than that obtained by activity screening using the wild-type protein as a template (6.7 %). When the highly active stability-maintaining variants were translated back to those derived from the previous round, the results were 33, 13, 30, 33, and 50 % in the order of the stability-maintaining variants derived from the first to fifth rounds. These values were also higher than that obtained in activity screening using the wild-type protein as a template. Thus, from the perspective of evolvability, stability screening using stability-maintaining variants was effective for obtaining highly active variants.

Using a fitness landscape analogy, the results suggest that the Aac-Est wild-type protein is located at a local peak that is surrounded by many downhill paths, and, that the improvement of enzyme activity by mutagenesis requires a climb. Fig. 5 shows an image of the inferred landscape of Aac-Est, which is expected to be similar to a rugged landscape with many peaks, such as Badlands National Park (South Dakota), as opposed to a smooth, single-peaked landscape, such as Mt. Fuji (Japan) [34,35]. If the Aac-Est wild-type protein is located near the top of a low mountain, downward and horizontal movements to different mountains are possible. When variant stability is maintained, the structure can be retained, even if it is destabilized by the addition of mutations. When the destination of migration while maintaining stability is at the foot of a higher mountain, this mountain can be climbed as mutations accumulate, indicating an increase in activity. Antibiotic resistance tests have used β-lactamase as a model enzyme to move among mountains by neutral selection or negative selection to find deleterious variants [36]; however, our study focused on stability indexes with Aac-Est. It has also been reported that within a cell, buffering the destabilizing effect of mutations with chaperones allows for a

Table 3
Characters of activity-enhancing variants obtained by activity screening after stability screenings.

Variants	Substitutions	Relative activity to WT at 65 °C
S2-A1	D130G F142L I149F T226S	1.4
S2-A2	P5H Q28L E50G T226S	1.3
S1-A1	D21Y K23T Y129F A268V	1.3
S6-A1	L25H V45E S185A	1.3
S4-S4	D52E A140E L222P S239R A249V	1.5
S2-S2-A1	V40A K42R K120R Q217H T226S I281V	1.3
S4-S2-S1-A1	Q33H K42R D52E L90I T295M	1.7
S4-S2-S1	K42R D52E L90I T295M	1.3
S4-S4-S1-A1	D52E A140E L222P S239R A249V K271E E273G	1.3
S4-S4-S1-S1-A1	H24L D52E D54G L100F A140E L222P S239R A249V K271E	1.5
S4-S4-S1-S3-A1	P39H D52E A136T A140E S221G L222P S239R A249V K271E	1.5
S4-S4-S1-S6	F51L D52E A140E L222P S239R A249V K271E I274V	1.4
S4-S4-S1-S2-A1	E47D D52E A140E L222P P236L S239R A249V K271E	1.4
S4-S4-S1-S3-A2	P39H D52E A140E S221G L222P S239R A249V V270D K271E	1.3
S4-S4-S1-S3-A3	P39H D52E A140E S221G L222P S239R A249V N266Y K271E	1.3
S4-S4-S1-S6-S2-A1	Q8R F51L D52E P72A A140E E199D L222P S239R A249V K271E I274V	1.6
S4-S4-S1-S6-S2-A2	Q8R F51L D52E P72A A140E N200K L222P S239R A249V K271E I274V	1.4
S4-S4-S1-S1-S1-A1	R16S D21G F51L D52E L100F A140E L222P S239R A249V K271E	1.4
S4-S4-S1-S3-S2-A1	P39H D52E A76V A140E I149T S221G L222P S239R A249V K267D K271E	1.4
S4-S4-S1-S6-S3-A1	A19S F51L D52E V87I A140E D141N L222P S239R A249V Y251C K271V K271V I274V R306Q	1.2

The underlined amino acid substitutions had already occurred in the template.

Table 4
Evolvability and highest relative activity at activity screening after stability screenings.

Templates	Evolvability (%)	Highest relative activity to WT at 65 °C
Aac-Est WT	6.7	1.2
Sx	27	1.4
Sx-Sx	13	1.5
Sx-Sx-Sx	30	1.7
Sx-Sx-Sx-Sx	40	1.5
Sx-Sx-Sx-Sx-Sx	50	1.6

protein to support genetic variation and evolution [37].

3.3. (Low-)stability and re-stabilization screenings

The majority of amino acid substitutions are accompanied by protein destabilization [12–14], and stabilization by amino acid substitutions is rare. As mutations accumulate, a protein becomes increasingly unstable and folding eventually becomes impossible [17]. Therefore, destabilization and the loss of folding ability intensify the decreased activity caused by amino acid substitutions. Variants with improved activity are difficult to obtain from mutations that use destabilized proteins as templates and these proteins are eliminated during the evolutionary process [16]. Previous studies have shown that it is challenging to improve the activity when the stability of the template protein is low

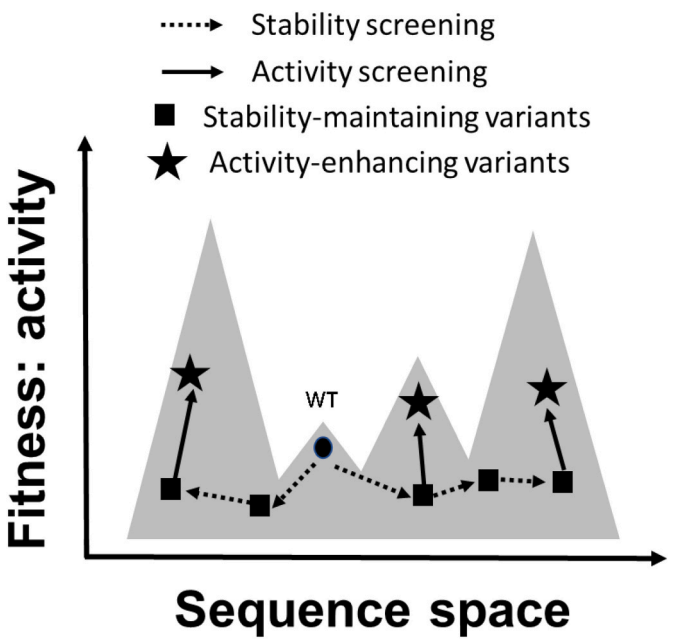


Fig. 5. Fitness landscape of the activity against the sequence space. Wild-type protein (circle) is located near the top of a low mountain. Stability-maintaining variants (squares) are obtained by stability screenings (dashed arrows). Activity-enhancing variants (stars) are obtained by activity-screenings (linear arrows).

[16,20,21]. Identifying stabilizing variants using mutations from destabilizing variants can lead to the discovery of a new evolutionary pathway.

During stability screening, variants that showed inadequate halo formation on plates at 65 °C but activity at 37 °C were termed "low-stabilized". Several low-stability variants were obtained from random mutations using wild-type protein and stability-maintaining variants as templates. Random mutations were then introduced into these low-

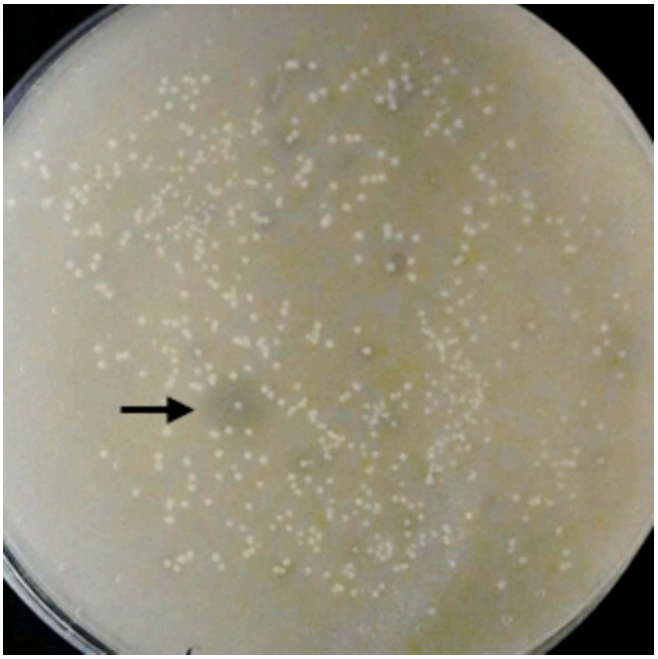


Fig. 6. An example of halo formation when a substrate-containing plate was layered directly on the transformed plate at 65 °C. The colony indicated by arrow forms a halo.

stabilizing variants to obtain restabilizing variants. Fig. 6 shows an example of halo formation when a substrate-containing plate was layered directly onto the transformed plate. A low-stability variant was used as the template protein; therefore, only a few halo-forming colonies were observed (estimated 1 halo colony out of >250). The method used in this study is effective for halo evaluations of a large number of colonies.

Table 5 shows the amino acid substitutions of the low-stabilizing variants and the eight restabilizing variants obtained. The halo activity evaluation using the 65 °C plate confirmed that the re-stabilizing variants had a greater halo. Of these, L2-R1, L3-R1, S2-L2-R1, S2-S2-L1-R1, and S2-S3-S1-L1-R1 remained active after heat treatment at 80 °C. The relative activities of these restabilizing variants relative to that of the wild-type at 65 °C are shown in Table 5. Many of the low-stability variants had low relative activities, whereas L2 showed a value of 0.90. This indicates a slight discrepancy between the halo activity and the activity measurement using the spectrometer, possibly a result of the destabilized L2 beginning to collapse at approximately 65 °C, thereby causing a difference in the results of the two measurement methods.

Although S2-S3-S1-L1-R1 showed low activity values, the activity of most restabilizing variants was equal to or higher than that of the wild-type protein. In particular, L2-R1 and S2-L2-R1 were hyperactive compared to the wild-type protein (relative activities of 1.5 and 1.4, respectively). These results indicate that restabilization screening enabled the creation of new variants that follow a new evolutionary pathway from that of the destabilized variants that were undergoing selection.

Among the restabilizing variants, the substitutions L115Q in L1-R1, Y65N in S4-L1-R1, and P253S in S2-S3-S1-L1-R1 were absent from the destabilizing variants. This implies that these mutations were destabilizing factors and that their elimination by reversion to the wild-type sequence led to restabilization. However, other variants introduced multiple mutations such as reversion to the wild-type sequence, further mutations at the mutation site, and new mutations, making it difficult to determine the effects of each. In addition, none of the mutants were restabilized with a single amino acid substitution. A protein engineering strategy has been reported in which stabilizing mutations are found by stabilizing a slightly destabilized parental protein [38]. However, in a previous study in which mutant libraries were generated in the background of introduced destabilizing variants of the bacterial toxin CcdB to identify stabilizing mutations, the effects of highly destabilizing mutations could not be altered by introducing point mutations [39]. Thus, restabilization may be more readily achieved by eliminating destabilizing factors as opposed to adding stabilizing mutations to compensate for instability. Therefore, stabilizing factors were examined for the highly activated L2-R1 and S2-L2-R1 variants.

Table 5
Characters of low-stability and re-stabilizing variants obtained by (low-)stability and re-stabilization screenings.

Variants	Substitutions	Relative activity to WT at 65 °C
L1	L115Q D141G K259E F287S	0.13
L1-R1	D141G K259E F287S	0.99
L2	K23N F51L M53V L180P K267Q K271E	0.90
L2-R1	K23N P44S F51L M53V F142L N266S K267Q K271E	1.5
L3	S32P D127G	0
L3-R1	S32P T92S Y235C	0.82
S2-L1	P39L T59M Y79H T226S L260P	0
S2-L1-R1	P39L T59M Y79H T226S L260R R300H	0.62
S4-L1	D52E Y65N P122T K267E E273G	0.54
S4-L1-R1	D52E P122T K267E E273G	1.1
S2-L2	G57D L167P T226S	0
S2-L2-R1	G57D L167S G208S T226S	1.4
S2-S2-L1	L11P K42R T59S K120R Q217H T226S I281V	0
S2-S2-L1-R1	K23N K42R T59S K120R A175D Q217H T226S I281V	1.1
S2-S3-S1-L1	F51L V70A V87D R138H L144I G203D T226S P253S A264E I274T	0
S2-S3-S1-L1-R1	F51L V70A V87D R138H L144I G203D T226S A264E I274T	0.22

The underlined amino acid substitutions indicate those that are unchanged in the low-stability and re-stabilizing variants.

The results for L2-R1 showed that L180P disappeared, while P44S, F142L, and N266S were introduced. Subsequently, L2-P180L was created by introducing P180L, a wild-type sequence revertant, into template L2. L2-P180L showed halo activity at 65 °C and 80 °C → 65 °C, with a relative activity of 1.1 at 65 °C. This suggests that L180P is a destabilizing factor, and its removal restored activity at 65 °C and the mutations in P44S, F142L, and N266S resulted in higher activity. In S2-L2-R1, L167P changed to S, and G208S was added, resulting in S2-L2-P167S. Similar to L2-P180L, S2-L2-P167S showed halo activity at 65 °C and 80 °C → 65 °C, with a relative activity of 1.1 at 65 °C. This suggests that L167P was a destabilizing factor and that its replacement with Ser eliminated destabilization and increased the activity of G208S. Therefore, mutations that eliminate destabilizing factors, rather than new stabilizing mutations that offset destabilization, appear to be effective for restabilization. Destabilized variants were restabilized to active variants at optimal temperatures by eliminating destabilizing factors, some of which could show more activity than the wild-type protein. This indicates the potential for low-stability variants with low-temperature activities.

Another notable mutation among the restabilizing variants was L167S in S2-L2-R1. In this case, the amino acid substitutions Leu → Pro→Ser occurred and two-base mutations in the nucleotide, CTT → CCT → TCT were required. Amino acid variations that can only be accessed by multiple base mutations in a codon are extremely rare when using methods that involve random mutations [40]. This suggests that low-stability variants may be useful to obtain amino acid substitutions that are not easily accessible.

4. Conclusions

Stability screening has led to the acquisition of more active variants from random mutant libraries using stability-maintaining variants as template proteins. In addition, the possibility of obtaining activity-enhancing variants is high, supporting previous studies that demonstrated the usefulness of the stability index.

In the restabilization of low-stability variants, the omission of replica plates enabled efficient evaluation of a large number of mutant libraries. Eight restabilized variants were obtained, some of which showed improved activities at optimal temperatures compared with the Aac-Est wild-type protein. Sequence analysis showed that the re-stabilizing variants did not contain additional mutations to compensate for destabilization, but rather reverted to the wild-type sequence at the mutation site or mutated at the same site. After restabilization to a thermally stable level, variants with increased activity were identified by the addition of further mutations. Therefore, low-stability variants can be re-evolved by removing destabilizing factors and these low-stability variants with low-temperature activity can be used. In addition, low-stability variants can be employed to identify amino acid substitutions that lead to the maintenance of more diverse stability and increased activity and reach amino acid substitutions that are not easily accessible,

requiring multiple base substitutions in codons. This study added a restabilization step to propose a new screening index to further diversify evolutionary pathways and evaluate variation-rich variants. The present study was performed on Aac-Est, but future studies on other proteins will establish the generality of this method.

CRediT authorship contribution statement

Mitsutoshi Wakisaka: Writing – original draft, Investigation. **Shunichi Tanaka:** Project administration, Funding acquisition. **Kazufumi Takano:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this study.

Acknowledgements

This work was supported in part by grants from the Japan Society for the Promotion of Science (KAKENHI No. 21K05386).

References

[1] R. Martínez, U. Schwaneberg, A roadmap to directed enzyme evolution and screening systems for biotechnological applications, Biol. Res. 46 (2013) 395–405.
[2] P.G. Chandler, S.S. Broendum, B.T. Riley, M.A. Spence, C.J. Jackson, S. McGowan, A.M. Buckle, Strategies for increasing protein stability, Methods Mol. Biol. 2073 (2020) 163–181.
[3] E.T. Farinas, T. Bulter, F.H. Arnold, Directed enzyme evolution, Curr. Opin. Biotechnol. 12 (2001) 545–551.
[4] R. Fasan, Y.T. Meharenna, C.D. Snow, T.L. Poulos, F.H. Arnold, Evolutionary history of a specialized P450 propane monooxygenase, J. Mol. Biol. 383 (2008) 1069–1080.
[5] M.S. Packer, D.R. Liu, Methods for the directed evolution of proteins, Nat. Rev. Genet. 16 (2015) 379–394.
[6] J.L. Porter, R.A. Rusli, D.L. Ollis, Directed evolution of enzymes for industrial biocatalysis, ChemBioChem 17 (2016) 197–203.
[7] S.D. Stimple, M.D. Smith, P.M. Tessier, Directed evolution methods for overcoming trade-offs between protein activity and stability, AICHE J. 66 (2020) e16814.
[8] K.L. Morley, R.J. Kazlauskas, R. J., Improving enzyme properties: when are closer mutations better? Trends Biotechnol. 23 (2005) 231–237.
[9] K. Chen, F.H. Arnold, Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide, Proc. Natl. Acad. Sci. USA 90 (1993) 5618–5622.
[10] A. Glieder, E.T. Farinas, F.H. Arnold, Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase, Nat. Biotechnol. 20 (2002) 1135–1139.
[11] P.A. Dalby, Strategy and success for the directed evolution of enzymes, Curr. Opin. Struct. Biol. 21 (2011) 473–480.
[12] S. Shafikhani, R.A. Siegel, E. Ferrari, V. Schellenberger, Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization, BioTechniques 23 (1997) 304–310.
[13] H.H. Guo, J. Choe, L.A. Loeb, Protein tolerance to random amino acid change, Proc. Natl. Acad. Sci. USA 101 (2004) 9205–9210.
[14] N. Tokuriki, F. Stricher, L. Serrano, D.S. Tawfik, How protein stability and new functions trade off, PLoS Comput. Biol. 4 (2008) e1000002.
[15] X. Wang, G. Minasov, B.K. Shoichet, Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs, J. Mol. Biol. 320 (2002) 85–95.
[16] R. Kurahashi, S.-I. Tanaka, K. Takano, Activity-stability trade-off in random mutant proteins, J. Biosci. Bioeng. 128 (2019) 405–409.
[17] J.D. Bloom, J.J. Silberg, C.O. Wilke, D.A. Drummond, C. Adami, F.H. Arnold, Thermodynamic prediction of protein neutrality, Proc. Natl. Acad. Sci. USA 102 (2005) 606–611.
[18] J.D. Bloom, S.T. Labthavikul, C.R. Otey, F.H. Arnold, Protein stability promotes evolvability, Proc. Natl. Acad. Sci. USA 103 (2006) 5869–5874.
[19] N. Tokuriki, D.S. Tawfik, Stability effects of mutations and protein evolvability, Curr. Opin. Struct. Biol. 19 (2009) 596–604.
[20] N. Ota, R. Kurahashi, S. Sano, K. Takano, The direction of protein evolution is destined by the stability, Biochimie 150 (2018) 100–109.
[21] R. Kurahashi, S. Sano, K. Takano, Protein evolution is potentially governed by protein stability: directed evolution of an esterase from the hyperthermophilic archaeon *Sulfolobus tokodaii*, J. Mol. Evol. 86 (2018) 283–292.
[22] R. Kurahashi, S.-I. Tanaka, K. Takano, Highly active enzymes produced by directed evolution with stability-based selection, Enzym. Microb. Technol. 140 (2020) 109626.
[23] G. Darland, T.D. Brock, *Bacillus acidocaldarius* sp. nov., an acidophilic thermophilic spore-forming bacterium, Microbiology 67 (1971) 9–15.
[24] K. Takano, A. Aoi, Y. Koga, S. Kanaya, Evolvability of thermophilic proteins from archaea and bacteria, Biochemistry 52 (2013) 4774–4780.
[25] H. Hemilä, T.T. Koivula, I. Palva, Hormone-sensitive lipase is closely related to several bacterial proteins, and distantly related to acetylcholinesterase and lipoprotein lipase: identification of a superfamily of esterases and lipases, Biochim. Biophys. Acta 1210 (1994) 249–253.
[26] G. Manco, E. Adinolfi, F.M. Pisani, G. Ottolina, G. Carrea, M. Rossi, Overexpression and properties of a new thermophilic and thermostable esterase from *Bacillus acidocaldarius* with sequence similarity to hormone-sensitive lipase subfamily, Biochem. J. 332 (1998) 203–212.
[27] F. Kovacic, A. Mandrysch, C. Poojari, B. Strodel, K.E. Jaeger, Structural features determining thermal adaptation of esterases, Protein Eng. Des. Sel. 29 (2016) 65–76.
[28] L. Mandrich, L. Merone, M. Pezzullo, L. Cipolla, F. Nicotra, M. Rossi, G. Manco, Role of the N terminus in enzyme activity, stability and specificity in thermophilic esterases belonging to the HSL family, J. Mol. Biol. 345 (2005) 501–512.
[29] H. Chahinian, Y.B. Ali, A. Abousalham, S. Petry, L. Mandrich, G. Manco, S. Canaan, L. Sarda, Substrate specificity and kinetic properties of enzymes belonging to the hormone-sensitive lipase family: comparison with non-lipolytic and lipolytic carboxylesterases, Biochim. Biophys. Acta 1738 (2005) 29–36.
[30] R.C. Cadwell, G.F. Joyce, Randomization of genes by PCR mutagenesis, PCR Methods Appl. 2 (1992) 28–33.
[31] T.S. Wong, D. Roccatano, M. Zacharias, U. Schwaneberg, A statistical analysis of random mutagenesis methods used for directed protein evolution, J. Mol. Biol. 355 (2006) 858–871.
[32] R.A. Beckman, A.S. Mildvan, L.A. Loeb, On the fidelity of DNA replication: manganese mutagenesis in vitro, Biochemistry 24 (1985) 5810–5817.
[33] S.-I. Tanaka, M. Tsutaki, S. Yamamoto, H. Mizutani, R. Kurahashi, A. Hirata, K. Takano, Exploring mutable conserved sites and fatal non-conserved sites by random mutation of esterase from *Sulfolobus tokodaii* and subtilisin from *Thermococcus kodakarensis*, Int. J. Biol. Macromol. 170 (2021) 343–353.
[34] F.J. Poelwijk, D.J. Kiviet, D.M. Weinreich, S.J. Tans, Empirical fitness landscapes reveal accessible evolutionary paths, Nature 445 (2007) 383–386.
[35] P.A. Romero, F.H. Arnold, Exploring protein fitness landscapes by directed evolution, Nat. Rev. Mol. Cell Biol. 10 (2009) 866–876.
[36] B. Steinberg, M. Ostermeier, Environmental changes bridge evolutionary valleys, Sci. Adv. 2 (2016) e1500921.
[37] N. Tokuriki, D.S. Tawfik, Chaperonin overexpression promotes genetic variation and enzyme evolution, Nature 459 (2009) 668–673.
[38] T. Tadokoro, K. Matsushita, Y. Abe, M.S. Rohman, Y. Koga, K. Takano, S. Kanaya, Remarkable stabilization of a psychrotrophic RNase HI by a combination of thermostabilizing mutations identified by the suppressor mutation method, Biochemistry 47 (2008) 8040–8047.
[39] S. Ahmed, K. Manjunath, G. Chattopadhyay, R. Varadarajan, Identification of stabilizing point mutations through mutagenesis of destabilized protein libraries, J. Biol. Chem. 298 (2022) 101785.
[40] K. Miyazaki, F.H. Arnold, Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function, J. Mol. Evol. 49 (1999) 716–720.