Artificial cysteine-lipases with high activity and altered catalytic mechanism created by laboratory evolution

Cen, Y., Singh, W., Arkin, M., Moody, T. S., Huang, M., Zhou, J., ... Reetz, M. T. (2019). Artificial cysteine-lipases with high activity and altered catalytic mechanism created by laboratory evolution. Nature Communications, 10, [3198]. https://doi.org/10.1038/s41467-019-11155-3

Published in:
Nature Communications

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2019 the authors.
This is an open access article published under a Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Artificial cysteine-lipases with high activity and altered catalytic mechanism created by laboratory evolution

Yixin Cen1,2, Warispreet Singh3,4, Mamatjan Arkin1, Thomas S. Moody4, Meilan Huang3, Jiahai Zhou2, Qi Wu1 & Manfred T. Reetz5,6

Engineering artificial enzymes with high activity and catalytic mechanism different from naturally occurring enzymes is a challenge in protein design. For example, many attempts have been made to obtain active hydrolases by introducing a Ser → Cys exchange at the respective catalytic triads, but this generally induced a breakdown of activity. We now report that this long-standing dogma no longer pertains, provided additional mutations are introduced by directed evolution. By employing Candida antarctica lipase B (CALB) as the model enzyme with the Ser-His-Asp catalytic triad, a highly active cysteine-lipase having a Cys-His-Asp catalytic triad and additional mutations W104V/A281Y/A282Y/V149G can be evolved, showing a 40-fold higher catalytic efficiency than wild-type CALB in the hydrolysis of 4-nitrophenyl benzoate, and tolerating bulky substrates. Crystal structures, kinetics, MD simulations and QM/MM calculations reveal dynamic features and explain all results, including the preference of a two-step mechanism involving the zwitterionic pair Cys105−/His224+ rather than a concerted process.
Lipases and proteases are textbook examples of hydrolases which employ a key set of active-site residues for hydrolyzing esters and amides, respectively, and constitute one of the largest enzyme families in the human proteome. Lipases function according to the classical Ser-His-Asp catalytic triad mechanism (Fig. 1a). Unlike lipases, proteases adopt divergent active sites and are grouped into seven mechanistic classes: serine-, cysteine- aspartic-, metallo-, threonine-, glutamic-, and asparagine-proteases. Among them, serine proteases have the Ser-His-Asp triad (Fig. 1b), while in the cysteine proteases, the catalytic triad consists of Cys-His-Asn (Fig. 1c). According to sequence alignment and phylogenetic analysis of proteases, the nucleophile exchanges Ser → Cys and Cys → Ser occurred during the evolution of both serine and cysteine proteases from common ancestors (Supplementary Fig. 1). Nevertheless, proteolytic active-site amino acids are the most evolutionarily conserved residues. The structural and mechanistic similarities between lipases and serine proteases inspired us to raise the fundamental question whether the conserved serine-lipase can be evolved into an active cysteine-lipase by directed evolution.

In serine and cysteine proteases, investigators have studied the interconversion of serine and cysteine by either chemical or recombinant technology. In the classical studies, it was originally thought that oxygen and sulfur possess similar chemical properties, and that the exchange of these nucleophile moieties would not influence catalysis notably. However, seminal experiments focusing on the generation of such mutant enzymes as Thiol-Subtilisin and Thiol-Trypsin demonstrated that these mutants are inactive toward common esters and amides, respectively. The reverse exchange in several cysteine-proteases, namely Cys → Ser, likewise causes notable reduction in activity (Supplementary Table 1). Similar results were found in the studies of other serine-hydrolases or cysteine-hydrolases (Supplementary Table 1). To the best of our knowledge, protein engineering of nucleophile-exchanged mutants with improved activity was never achieved.

For lipases, only a few reports of cysteine analogs have been reported, all resulting in greatly reduced enzyme activity (Supplementary Table 1). In contrast to native proteases, which possess either serine or cysteine nucleophilic residues, lipases are highly conserved with a serine nucleophile. It is therefore methodologically more challenging to engineer serine-lipases into highly active cysteine-lipases. In the present study we propose that the Ser → Cys conversion in lipases may lead to high activity by accurate manipulation of the local environment surrounding the hybrid Cys-His-Asp triad, enabled by laboratory evolution (Fig. 1d). We not only demonstrate this experimentally, but also provide an up to date theoretical analysis why the Ser to Cys exchange causes extreme activity reduction, while additional mutations result in high activity when testing substrates which are essentially not accepted by wild-type (WT).

Here, we choose lipase B from Candida antarctica (CALB), as the model enzyme, in which Asp187-His224-Ser105 is the catalytically active triad. As expected, Thiol-CALB (variant QW1; S105C) proves to be essentially inactive in the hydrolysis of such substrates as 1 and 4–19 (Fig. 2c). We succeed in evolving cysteine-CALB mutants that are even more active than WT CALB. Structural, kinetic and theoretical investigations point to a distinct catalytic mechanism, different from all serine-lipases known to date.

### Results

#### Directed evolution of cysteine-CALB

In order to examine the influence of the nucleophile exchange Ser → Cys on CALB activity, we selected ester 1 as the model substrate, which is hardly accepted by WT due to steric hindrance of the bulky benzyl group. Kinetic experiments using purified variant S105C (QW1) show that this mutation diminishes the catalytic efficiency by a factor of about two, with $k_{cat}/K_m$ (WT CALB) being $150 \text{ s}^{-1}\text{M}^{-1}$ compared to $k_{cat}/K_m = 88 \text{ s}^{-1}\text{M}^{-1}$ for mutant QW1 (Table 1).

Protein engineering was pursued in order to achieve high activity of cysteine-CALB. Directed evolution and rational site-specific mutagenesis of enzymes are well known tools for improving their catalytic properties. We used iterative saturation mutagenesis (ISM) at sites lining the binding pocket, a well established technique. The choice of randomization sites was guided by the crystal structure of WT CALB (PDB code...
Fig. 2 The choice of the ISM sites of CALB and enzymatic transformations of different substrates. a Selected 2-residue randomization sites A-F for ISM based on the X-ray structure of CALB. Site A (W104/S105, red), B (A281/A282, green), C (L144/V149, blue), D (I189/V190, yellow), E (L140/A141, purple), and F (L278/I285, cyan). The catalytic triad Asp187-His224-Ser105 are shown in gray and red stick illustrations. b Binding pocket surface of WT CALB with the docked substrate 1. Mutagenesis sites A, B, C, D, E and F are shown as surface illustration with the same color coding as in a. c–d, Structure and reactions of tested substrates.
Table 1 Kinetic data of WT CALB and mutants

| Substrate | Entry | Enzymes | Mutations | $k_{cat}$ (s$^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ (s$^{-1}$M$^{-1}$) |
|-----------|-------|---------|-----------|---------------------|-----------|---------------------|
| 1 | WT | - | - | 0.025 ± 0.009 | 0.17 ± 0.08 | 150 ± 11 |
| 2 | QW1 | S105C | - | 0.028 ± 0.007 | 0.32 ± 0.02 | 88 ± 17 |
| 3 | QW2 | W104V/S105C | - | 0.032 ± 0.005 | 0.50 ± 0.12 | 64 ± 6 |
| 4 | QW3 | W104V/S105C/A281Y/A282Y | - | 0.031 ± 0.008 | 0.051 ± 0.02 | 610 ± 100 |
| 5 | QW4 | W104V/S105C/A281Y/A282Y/V149G | - | 0.187 ± 0.024 | 0.032 ± 0.01 | 5920 ± 340 |
| 6 | QW5 | W104V | - | 0.016 ± 0.003 | 0.014 ± 0.004 | 1187 ± 174 |
| 7 | QW6 | V149G | - | 0.059 ± 0.009 | 0.077 ± 0.015 | 772 ± 48 |
| 8 | QW7 | A281Y | - | 0.043 ± 0.011 | 0.066 ± 0.013 | 653 ± 100 |
| 9 | QW8 | A282Y | - | 0.052 ± 0.014 | 0.513 ± 0.083 | 98 ± 16 |
| 10 | QW9 | S105C/A281Y/A282Y/V149G | - | 0.041 ± 0.008 | 0.066 ± 0.011 | 610 ± 97 |
| 11 | QW10 | W104V/A281Y/A282Y/V149G | - | 0.211 ± 0.012 | 0.43 ± 0.001 | 4914 ± 251 |
| 12 | QW11 | W104V/S105C/A281Y/A282Y/V149G | - | 0.027 ± 0.003 | 0.015 ± 0.004 | 1843 ± 313 |
| 13 | QW12 | W104V/S105C/A281Y/V149G | - | 0.071 ± 0.006 | 0.031 ± 0.001 | 2256 ± 198 |
| 14 | WT | - | - | 0.030 ± 0.012 | 0.197 ± 0.093 | 150 ± 10 |
| 15 | QW1 | S105C | - | 0.060 ± 0.003 | 0.134 ± 0.012 | 450 ± 18 |
| 16 | QW4 | W104V/S105C/A281Y/A282Y/V149G | - | 1.605 ± 0.28 | 0.246 ± 0.059 | 6610 ± 510 |
| 17 | QW10 | W104V/A281Y/A282Y/V149G | - | 2.609 ± 0.28 | 0.565 ± 0.070 | 4620 ± 70 |
| 18 | WT | - | - | 0.013 ± 0.001 | 0.018 ± 0.003 | 750 ± 140 |
| 19 | QW1 | S105C | - | 0.008 ± 0.001 | 0.012 ± 0.003 | 660 ± 85 |
| 20 | QW4 | W104V/S105C/A281Y/A282Y/V149G | - | 0.460 ± 0.026 | 0.075 ± 0.010 | 6140 ± 362 |
| 21 | QW10 | W104V/A281Y/A282Y/V149G | - | 3.548 ± 0.583 | 0.424 ± 0.120 | 8426 ± 276 |
| 22 | WT | - | - | 0.022 ± 0.003 | 0.024 ± 0.018 | 1120 ± 320 |
| 23 | QW1 | S105C | - | 0.042 ± 0.019 | 0.044 ± 0.025 | 990 ± 140 |
| 24 | QW4 | W104V/S105C/A281Y/A282Y/V149G | - | 0.903 ± 0.179 | 0.138 ± 0.033 | 6550 ± 290 |
| 25 | QW10 | W104V/A281Y/A282Y/V149G | - | 4.15 ± 0.96 | 0.630 ± 0.153 | 6600 ± 93 |
| 26 | WT | - | - | 0.099 ± 0.013 | 0.188 ± 0.021 | 523 ± 11 |
| 27 | QW1 | S105C | - | 0.031 ± 0.004 | 0.075 ± 0.004 | 459 ± 25 |
| 28 | QW4 | W104V/S105C/A281Y/A282Y/V149G | - | 0.165 ± 0.007 | 0.039 ± 0.008 | 4310 ± 346 |
| 29 | QW10 | W104V/A281Y/A282Y/V149G | - | 0.307 ± 0.050 | 0.114 ± 0.014 | 2675 ± 115 |

Note: Source data are provided as a Source Data file.
QW3, and QW11) when compared with the best variant QW4 (Table 1, entries 10, 4, 12). Activity assessments of these single point variants and multiple-point variants indicate that the effect of these mutations is cooperative rather than additive.

Notably, variant QW10, which has the natural catalytic triad Ser-His-Asp restored and only differs from the best variant QW4 by the residue at position 105, displays lower activity than QW4 with the hybrid Cys-His-Asp triad (Table 1, entry 11). This clearly implies the feasibility of creating a highly active cysteine-lipase with the hybrid Cys-His-Asp triad by means of directed evolution.

Substrate scope. A series of other bulky esters were also tested as expanded substrates, because most of them are not well accepted by WT CALB due to steric hindrance (Table 1, entries 14–29). The kinetic data shows, inter alia, that variants QW4 and QW10 are much more active for several substituted benzoates than WT and QW1. When comparing kinetic data or reaction time course, QW4 is more efficient than QW10 in the conversion of 4 and 7, and almost as active as QW10 for 5 and 6 (Table 1, Supplementary Fig. 3). Other sterically hindered substrates such as 

\[ \text{p-nitrophenyl cyclohexanecarboxylate} \]  

or 

\[ \text{p-nitrophenyl cycloheptanecarboxylate} \]  

are also efficiently hydrolyzed by QW4 and QW10 (Supplementary Fig. 3f–g).

WT CALB is an excellent catalyst in the hydrolytic kinetic resolution (KR) of appropriate esters of chiral secondary alcohols with preferential formation of (R)-enantiomers\(^2\). The present mutants were tested in the reactions of four sec-alcohol esters, \(\text{rac-11} \sim \text{rac-17}\). As expected, WT showed good selectivity and activity for \(\text{rac-11} (E > 200 \text{ (R)}, \text{Supplementary Table 3, entry 10, 4, 12})\) and \(\text{rac-15} (E = 83 \text{ (R)}, \text{Table 2, entry 1})\), while QW1 showed reduced activity and low enantioselectivity. Upon using the best mutant QW4, notably improved conversion was observed compared with QW1 (Table 2, Supplementary Table 3). Remarkably, Q4 induces reversed enantioselectivity in favor of the (S)-alcohol, similar to variant W104A previously reported by Hult et al.\(^2\). In the case of \(\text{rac-15}\), variant QW4 is superior to QW10 (Table 2), and for \(\text{rac-13}\) they are similar, while for \(\text{rac-11}\) QW10 is a little better (Supplementary Table 3). It is noteworthy that \(\text{rac-17}\) with a large alkyl group also can be accepted by QW4 and QW10 with moderate selectivity, in sharp contrast to WT CALB which has no activity toward this substrate due to the limited space of the alcohol-binding pocket (Supplementary Table 3, entry 11–14). Moreover, WT CALB is also a poor catalyst in the KR of \(\text{rac-19}\) with the stereogenic center in the acid fragment, and low activity and poor enantioselectivity were observed \((E = 2 \text{ (R)}, \text{Supplementary Table 3, entry 15})\). Thiol-CALB and other mutants also perform poorly for this type of esters (Supplementary Table 3, entry 16–18), which outlines the limitation of the CALB mutants.

In the case of KR of \(\text{trans-21}\) (trans-dimethyl cyclohexane-1,2-dicarboxylate), surprisingly, QW1 (Thiol-CALB) displays unexpectedly excellent performance for this specific substrate like WT CALB \((E > 200 \text{ (R)}, \text{Supplementary Table 3, entry 19–20})\). However, both QW4 and QW10 variants showed very low activity for \(\text{trans-21}\), probably because their reshaped active sites cannot accept this substrate well. Interestingly, high stereoselectivity for \(\text{trans-21}\) with ee values up to 99% was also observed for this reaction catalyzed by QW4 and QW10 variants, respectively (Supplementary Table 3, entry 21–22).

We further tested the nonaqueous transesterification of the model substrate (1) using methanol catalysed by WT CALB and QW4, QW10 variants. The results clearly demonstrate excellent performance of the cysteine-lipase QW4 for this bulky substrate, in distinct contrast to WT CALB (Supplementary Fig. 4).

Crystal structure characterization. In order to gain insight into the basis of increased activities of QW4 and QW10, the crystal structures of both enzymes were solved and refined to 1.85 and 1.88 Å resolution, respectively. Surprisingly, Cys105 in variant QW4 appeared to be oxidized to the respective sulfonic acid. When purification and crystallization were performed in a glove box, Cys105 in the crystal structure was in the normal (reduced) state. In fact, cysteine proteases are readily oxidized\(^2\) to give the disulfide bond, Cys-Cys, in the absence of thiol reductant. A wide range of cysteine proteases is reported to be autocatalytically oxidized to the disulfide bond to give inactive, non-catalytic thiol proteases, which may imply that the active site sulfur in QW4 is reactive.

Most residues in the structures of QW4 and QW10 have very similar conformations. The most remarkable difference was observed at loop 137–150 and helix 277–288 (α-10 helix\(^2\)). Loop 137–150 of QW4 and QW10, the lid at the entrance of catalytic sites in CALB\(^2\), is very flexible as compared to those in WT. This was indicated by higher B-factors (Fig. 3a–c) and RMSF in subsequent molecular dynamics (MD) simulation (Supplementary Fig. 8). The raised mobility of loop 137–150 in QW4 and QW10 enables ready entry of bulky substrates\(^2\), which can be attributed to the introduction of the V149G mutation. In addition, there are different H-bond networks between the α-6 helix (residues 151–157), loop 137–150 and α-10 helix (residues 267 to 288) in QW4, QW10, and WT-CALB (Fig. 3g–i). In QW10 and WT, three or four H-bonds exist between S150/W155 and Q291 (Fig. 3g, i), while in QW4 there are only two H-bonds between G149/W155 and Q291 (Fig. 3h). These differences in the H-bond interactions result in a more flexible and wider open lid in QW4 than those in QW10 and WT (Fig. 3h, e).

In addition, in the crystal structure of QW4, the B-factors of the α-10 helix (residues 267–288), especially at helix 277–288, are much higher than those in WT and QW10 (Fig. 3a versus 3b–c). This increased flexibility may aid the substrate to enter the binding pocket and to bind at the active site, thus leading to the high activity of QW4 in the hydrolysis of substrate 1. Interestingly, the nucleophile exchange (S105C) in QW4 results in a conformational change of helix 277–288, compared with...
of the substrate into the proximity of the active site and results in a bent conformation of 1 (Fig. 4a, Supplementary Figs. 7a and 9).

For QW10 and QW4 (Fig. 4c, d), S105 and C105 remain in close proximity to the substrate carbonyl group, favorable for nucleophilic attack to occur (Supplementary Fig. 7d-e). Compared with QW2 (Fig. 4b), introducing the A281Y/A282Y mutations in QW4 and QW10 causes obvious displacement around the helix region 279–289. P280 and Y281 move in the proximity to the substrate, and form favorable hydrophobic stacking interaction with the phenyl ring of 1. The loop following the helix 279–289 moves accordingly, so that the H-bond between K290 and V286 observed in WT and QW2 is broken, and a new H-bond interaction is established between G149 and Q291, resulting in an open conformation of the lid. Extra space is created by the induced open conformation of the QW4 and QW10 variants (Fig. 4e, f), which may also account for the high activities of the best mutants. The interactions with Y281 and E188 in QW4 further positions the substrate in the catalytic pocket of the enzyme. Moreover, the presence of a water molecule, which forms an

Substrate binding disclosed by MD simulations. In order to interpret more precisely the differences in catalytic activities of these enzymes in presence of a substrate, p-nitrophenyl benzoate (1) was docked into the binding pockets of WT-CALB, and variants QW2, QW4, and QW10, respectively, and 100-ns MD simulations were run. The results revealed distinctly different modes of binding in WT and in variants QW4 and QW10 (Fig. 4). It was observed that W104 in WT clearly hinders access to

Fig. 3 Cartoon representation of the crystal structures of enzymes. a Variant QW4. b Variant QW10. c WT CALB (PDB: 5a71). a-c are colored according to the B-factor. d Superimposed overall structures of WT CALB (gray), QW4 (green) and QW10 (pink). e Zoom-in view of the loop 137–150 in WT CALB (gray), in QW4 (green) and in QW10 (pink). f Zoom-in view of helix 277–288 in WT CALB (gray), in QW4 (green) and in QW10 (pink). Hydrogen bonds formed between the α-6 helix, loop 137–150 and α-10 helix in g (WT-CALB), h (QW4) and i (QW10). H-bond formed between Tyr281 and Glu188 in QW4 mutant (h) was also noted.

QW10 or WT. While the helixes of QW10 and WT CALB are well superimposable (Fig. 3f), the helix section P280-Y282 of QW4 is deformed and shifted toward the active site, with an additional H-bond between Y281 and E188 formed (Fig. 3h). This conformational change also induces loop 289–294 next to helix 277–288 to shift by about 3.8 Å towards the solvent. This effect leads to more room for the substrate at the entrance to the binding pocket (Fig. 3f). By comparing the kinetic data of QW3 and QW2, we found that $K_m$ is reduced by a factor of 10 (Table 1, entries 3–4). This indicates that the above noted movement of helix 277–288 is likely to be present in QW4, where the A281Y/A282Y mutation is also introduced.
H-bond with substrate 1 in QW4 (Fig. 4d), may be critical for the subsequent hydrolysis reaction, and therefore also be responsible for its high catalytic efficiency.

**Altered reaction mechanism of cysteine-CALB variant QW4.** In order to compare the reaction processes and mechanistic details of lipases with naturally occurring Ser-His-Asp triad versus engineered Cys-His-Asp triad on a molecular level, we applied QM/MM studies on cysteine proteases, including human Cathepsin K and papain featuring a Cys-His-Asp triad, indicated that the catalytic Cys remains protonated during nucleophilic attack in the formation process of acyl-enzyme complexes. Previous QM/MM studies on cysteine proteases of the step-wise alternative.

In striking contrast, the potential energy surface scan for QW4 disclosed that the acylation reaction follows a two-step mechanism involving a HisH+ /CysS− ion pair, in which the proton is first transferred from the thiol group of C105 to H224, and then the negatively charged thiolate anion acts as a nucleophile, attacking the carbon of ester bond of the substrate to yield the tetrahedral intermediate. The first step, i.e., the proton transfer from SH of C105 to H224 is the rate determining step with an activation barrier of 13.95 kcal/mol, and then the thiolate anion attacks the carbonyl of the substrate rapidly with a low barrier of 6.2 kcal/mol (Fig. 5b, Supplementary Fig. 12).

In further work, we conducted QM cluster calculations, hoping to validate the two-step reaction mechanism proposed for variant QW4. Starting from the ionic pair, the estimated energy barrier and transition state are in agreement with the QM/MM results.
The concerted mechanism was also examined for QW4 using the QM cluster model. Again, no transition state corresponding to a concerted mechanism was located (Supplementary Fig. 15).

Discussion

It is known that the overall topologies and catalytic properties of most serine and cysteine hydrolases are remarkably different, indicating that they evolved separately as a result of convergent evolution. On the other hand, nucleophile exchanges of this kind do occur in the evolutionary history of serine and cysteine proteases. For example, viral cysteine proteases appear to be homologous to the trypsin family of eukaryotic serine proteases. These cases indicate that the interconversion of nucleophilic serine and cysteine cause notable effects of the normal CALB variant QW10 and cysteine-CALB variant QW4. The concerted mechanism of simultaneous proton transfer and nucleophilic attack. For the cysteine-CALB variant QW4, the formation of the tetrahedral intermediate occurs via a two-step reaction mechanism where proton transfer from Cys105 to His224 is followed by a nucleophilic attack of the deprotonated Cys105 to the carbonyl group of the substrate. The reaction profiles in the hydrolysis of substrate 1 were calculated by QM/MM (AMBER/ωB97X-D/6-31G+(d,p)). The reaction coordinate is the distance between the substrate carbonyl carbon and hydroxyl oxygen of Ser105 or thiolate sulfur of deprotonated Cys105.

Modern QM/MM techniques not only allow for the elucidation of mechanistic details of enzymes, they can also be used to predict enzyme reaction mechanisms. In contrast to the widely studied reaction mechanism of proteases, little information is available concerning the mechanism of lipases based on QM/MM calculations. On the basis of X-ray structural data and QM/MM calculations, we observed the difference in reaction mechanism between CALB with the naturally occurring triad of Ser-His-Asp and the cysteine-CALB mutant (QW4) with a Cys-His-Asp catalytic triad, which actually enables otherwise difficult transformations.

The concerted mechanism of proteases is also known to be a mechanism that may be accelerated by a conformational change in the enzyme. This change allows the enzyme to adopt a conformation that favors the nucleophilic attack of the cysteine residue. The concerted mechanism was also examined for QW4 using the QM cluster model. Again, no transition state corresponding to a concerted mechanism was located (Supplementary Fig. 15).

The concerted mechanism was also examined for QW4 using the QM cluster model. Again, no transition state corresponding to a concerted mechanism was located (Supplementary Fig. 15).
to its optimal binding properties, as well as the relatively low reaction barrier in the rate-determining step.

In summary, the successful construction of a highly active cysteine-lipase mutant with hybrid Cys-His-Asp triad expands the types of lipases that can be discovered and be functionally active. Our results suggest that besides the protease enzyme family, the occurrence of a Ser to Cys nucleophile exchange in the lipase family is also possible, in this case by laboratory evolution. This raises the intriguing question whether cysteine-lipases can be found in nature. The present work provides a valuable insight into the catalytic mechanisms of naturally occurring lipases and artificial cysteine-lipases at the molecular level and how these catalysts can transform structurally different substrates which are not accepted by the WT lipase. This opens up exciting new opportunities for these engineered lipases in chemical synthesis.

**Methods**

Library screening and kinetic measurements. Saturation mutagenesis libraries were constructed at sites A (Trp104/Ser105), B (Ala281/Ala282), C (Leu144/Val149), D (Ile189/Val190), E (Leu140/Ala141), and F (Leu278/Ile285). PCRAs were performed using WT-CALB plasmid (pETM11-CALB) as the template DNA, and forward primers (see Supplementary Table 5) and a silent reverse primer (GGGAGCAGACAAGCGGCGTACGAGG4 2444–2466 bp of pETM11). The reaction (100 μl final volume) contained: 10 x KOD buffer (10 μl), MgCl₂ (4 μl, 25 mM), dNTP (10 μl, 2 mM each), forward primers (4 μl, 2.5 μM each), silent reverse primer (4 μl, 2.5 μM each), template plasmid (1 μl, 100 ng μl⁻¹) and 1 μl of KOD polymerase. PCR conditions used were 95 °C, 3 min; five cycles of (98 °C, 1 min; 65 °C, 1 min; 72 °C, 1 min); and final extension at 72 °C, 10 min. The purified PCR products were transformed into electroporated cells of E. coli BL21 (DE3) (containing chaperone plasmid pG707, Takara, Japan). Transformants grown on LB-agar plates containing Kanamycin (34 μg/ml) and chloramphenicol (34 μg/ml) were picked up and cultured overnight at 37 °C with shaking (800 rpm) in 96 deep well plates containing 800 μl TB with Kanamycin (34 μg/ml) and chloramphenicol (34 μg/ml). After inoculation into fresh TB media containing antibiotics and L-arabinose (1 mg/ml) for 4 h at 37 °C, isoamyl β-thiogalactopyranoside (IPTG) (final concentration 1 mM) was added to induce the expression of CALB. After expression for 24 h at 18 °C, cells were harvested by centrifugation at 2750 x g and 4 °C for 25 min., and treated with lysosome and Dnase I. The supernatants of CALB libraries were transferred to microtiter plates for screening using a UV/Vis-plate reader. The kinetics data of purified enzymes as catalysts were determined on a UV/Vis plate reader by monitoring the time-dependent appearance of p-nitrophenolate (3) in the hydrolysis reactions of substrate 1, 4–7 at various concentration ranges. The obtained data were fitted to the Michaelis-Menten equation by nonlinear regression analysis.

**Data availability**

The data that support the findings of this study are available from the corresponding author upon reasonable request. The crystal structure of mutants QW4 (oxidized), QW4 (unoxidized) and QW10 has been deposited in the Protein Data Bank (PDB) under accession code 6ISQ, 6ISR, and 6ISP, respectively. The raw data underlying Table 1 and Supplementary Figures 2–4, 11, 13, 15 and Supplementary Table 2 are provided as a Source Data file.

Received: 18 December 2018 Accepted: 24 June 2019

**References**

1. Botos, I. & Wlodawer, A. The expanding diversity of serine hydrolases. *Curr. Opin. Struct. Biol.* 17, 683–690 (2007).
2. Schmid, R. D. & Verger, R. Lipases: Interfacial enzymes with attractive applications. *Angew. Chem. Int. Ed.* 37, 1608–1633 (1998).
3. Oda, K. New families of carboxyl peptides: serine-carboxyl peptides and glutamic peptides. *J. Biochem.* 151, 13–25 (2012).
4. Hedstrom, L. Serine protease mechanism and specificity. *Chem. Rev.* 102, 4501–4524 (2002).
5. Ekici, O. D., Paetzel, M. & Dalbey, R. E. Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* 17, 2023–2037 (2008).
6. Carter, P. & Wells, J. A. Dissecting the catalytic triad of a serine protease. *Nature 332*, 564–568 (1988).
7. Corey, D. R., McGrath, E. M., Vasquez, J. R., Fletterick, R. J. & Craik, C. S. An alternate geometry for the catalytic triad of serine proteases. *J. Am. Chem. Soc.* 114, 4905–4907 (1992).
8. Rawlings, N. D., Barrett, A. J. & Bateman, A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 40, D343–D350 (2012).
9. Neet, K. E. & Koshold, D. E. Jr. The conversion of serine at the active site of subtilisin to cysteine: A “chemical mutation”. *Proc. Natl Acad. Sci. USA* 56, 1606–1611 (1966).
10. Polgar, L. & Bender, M. L. A new enzyme containing a synthetically formed active site Thiol-subtilisin. *J. Am. Chem. Soc.* 88, 3153–3154 (1966).
11. Higaki, J. N., Envin, L. B. & Craik, C. S. Introduction of a cysteine protease active site into trypsin. *Biochemistry* 28, 9256–9263 (1989).
12. Slade, A., Horrocks, A. J., Lindsdell, C. J., Dunbar, B. & Virdee, R. Site-directed chemical conversion of serine to cysteine in penicillin acylase from *Escherichia coli* ATCC 11105: Effect on conformation and catalytic activity. *Eur. J. Biochem.* 175, 79–85 (1990).
13. Wilke, M. E., Higaki, J. N., Craik, C. S. & Fletterick, R. J. Crystal structure of rat tryptase Ser195C at −150 °C: Analysis of low activity of recombinant and prothrombin variant synthetic thiol proteases. *J. Mol. Biol.* 219, 511–523 (1991).
14. Hahn, C. H., Strauss, J. H. Site-directed mutagenesis of the proposed catalytic amino acids of the sindbis virus capsid protein autoprotease. *J. Virol.* 64, 3069–3073 (1990).
15. Bornscheuer, U. T. & Kalaivas, R. J. Hydrolyases in Organic Synthesis: Regio- and Stereoselective Biotransformations. (Wiley-VCH Weinheim, Germany, 1999).
16. Anderson, E. M., Larsson, K. M. & Kirk, O. One biocatalyst—many applications: the use of Candida antarctica B-lipase in organic synthesis. *Biocatal* 16, 181–204 (1998).
17. Qu, G., Li, A., Aceredo-Rocha, C. G., Sun, Z. & Reetz, M. T. The crucial role of methodology development in directed evolution of selective enzymes. *Angew. Chem. Int. Ed.* 58, https://doi.org/10.1002/anie.201901491 (2019).
18. Hammer, S. C., Knight, A. M. & Arnold, F. H. Design and evolution of enzymes for non-natural chemistry. *Curr. Opin. Green. Sustain. Chem.* 7, 23–30 (2017).
19. Zeeman, C. & Hilvert, D. Directed evolution of protein catalysts. *Annu. Rev. Biochem.* 87, 131–157 (2018).
20. Uppenberg, J., Hansen, M. T., Patkar, S. & Jones, T. A. The sequence, crystal structure determination and refinement of two crystal forms of lipase B from Candida antarctica. *Structure* 2, 293–308 (1994).
21. Stauth, B., Fisher, S. J. & Cianci, M. Open and closed states of Candida antarctica lipase B: Confirmation and the mechanism of interfacial activation. *J. Lipid Res.* 56, 2348–2358 (2015).
22. Magnusson, A. O., Takwa, M., Hamborg, A. & Hult, K. An S-selective lipase was regularly developed by rational redesign and the enantioselectivity increased with temperature. *Angew. Chem. Int. Ed.* 117, 4658–4661 (2005).
23. Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A. & Drenth, J. Structure of papain refined at 1.65 Å resolution. *J. Biol. Chem.* 179, 233–256 (1948).
24. Pathak, D., Ashley, G. & Ollis, D. Thiol protease-like active site found in the enzyme dienelactone hydrolase: localization using biochemical, genetic, and structural tools. *Proteins* 9, 267–279 (1991).

**Reporting summary**

Further information on research design is available in the Nature Research Reporting Summary linked to this article.
25. Kulaithu, Y. et al. Regulation of A20 and other OUT deubiquitinases by reversible oxidation. Nat. Commun. 4, 1569 (2013).

26. Waridel, A. Energetics of enzyme catalysis. Proc. Natl Acad. Sci. USA 75, 5250–5254 (1978).

27. McGrath, M. E., Wilke, M. E., Higaki, J. N., Craik, C. S. & Fletterick, R. J. Crystal structures of two engineered thioph trypsins. Biochemistry 28, 9264–9270 (1989).

28. Dorset et al. Structure and function of alternative proton-relay mutants of dihydrofolate reductase. Biochemistry 31, 9813–9822 (1992).

29. Skjøt, M. et al. Understanding the plasticity of the alpha/beta hydrolase fold: lid swapping on the Candida antarctica lipase B results in chimeras with interesting biocatalytic properties. Chem. Bio. Chem. 10, 520–527 (2009).

30. Gruber, C. C. & Pleiss, J. Lipase B from Candida antarctica binds to hydrophobic substrate-water interfaces via hydrophobic anchors surrounding the active site entrance. J. Mol. Catal. B 84, 48–54 (2012).

31. Zisis, T. et al. Interfacial activation of Candida antarctica lipase B: Combined evidence from experiment and simulation. Biochemistry 54, 5969–5979 (2015).

32. Senn, M. J. & Thiel, W. QM/MM methods for biomolecular systems. Angew. Chem. Int. Ed. 48, 1198–1229 (2009).

33. van der Kemp, M. W. & Mulholland, A. J. Combined quantum mechanical/molecular mechanics (QM/MM) methods in computational enzymology. Biochemistry 52, 2708–2728 (2013).

34. Ma, S., Devi-Kesavan, L. S. & Gao, J. Molecular dynamics simulations of the catalytic pathway of a cysteine protease: a combined QM/MM study of human cathepsin K. J. Am. Chem. Soc. 129, 13633–13645 (2007).

35. Wei, D., Huang, X., Liu, J., Tang, M. & Zhan, C. G. Reaction pathway and free energy profile for papain-catalyzed hydrolysis of N-Acetyl-Phe-Gly-4-Nitroanilide. Biochemistry 52, 5145–5154 (2013).

36. Elasser, B. et al. Distinct roles of catalytic cysteine and histidine in the proteolytic and ligase mechanisms of human legumain as revealed by DFT-based QM/MM simulations. ACS Catal. 7, 5585–5593 (2017).

37. Uritsky, N., Shokhen, M. & Albeck, A. Stepwise versus concerted mechanisms in general-base catalysis by serine proteases. Angew. Chem. Int. Ed. 55, 1680–1684 (2016).

38. Ishida, T. & Kato, S. Theoretical perspectives on the reaction mechanism of serine proteases: the reaction free energy profiles of the acylation process. J. Am. Chem. Soc. 125, 12035–12048 (2003).

39. Pinto, G. P. et al. New insights in the catalytic mechanism of tyrosine ammonia-lyase given by QM/MM and QM cluster models. Arch. Biochem. Biophys. 582, 107–115 (2015).

40. Bazan, J. F. & Fletterick, R. J. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. Proc. Natl Acad. Sci. USA 85, 7872–7876 (1988).

41. Sárkány, Z. & Polgár, L. The unusual catalytic triad of poliovirus protease 3C. Biophys. Chem. 42, 516–522 (2003).

42. Świderek, K., Marti, S. & Moliner, V. Theoretical study of primary reaction of pseudozyma antarctica lipase B as the starting point to understand its promiscuity. ACS Catal. 4, 426–434 (2014).

43. Escorcia, A. M., Sen, K., Daza, M. C., Doerr, M. & Thiel, W. Quantum mechanics/molecular mechanics insights into the enantioselectivity of the O-acetylation of (R,S)-propranolol catalyzed by Candida antarctica lipase B. ACS Catal. 7, 119–127 (2017).

44. Sanchis, J. et al. Improved PCR method for the creation of saturation mutagenesis libraries in directed evolution: application to difficult-to-amplify templates. Appl. Microbiol. Biotechnol. 81, 387–397 (2008).

45. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in general-case catalysis by serine proteases: the reaction free energy profile for papain-catalyzed hydrolysis of N-Acetyl-Phe-Gly-4-Nitroanilide. Biochemistry 52, 5145–5154 (2013).

46. Sanchis, J. et al. Improved PCR method for the creation of saturation mutagenesis libraries in directed evolution: application to difficult-to-amplify templates. Appl. Microbiol. Biotechnol. 81, 387–397 (2008).

47. Kabsch, W. Integration, scaling, space-group assignment and post-reduction. Acta Crystallogr. D Biol. Crystallogr. 66, 133–144 (2010).

48. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).