Publishable summary report

Proposal No. 236392

Integrated protein evolution model

The aim of the project was to identify epistatic effects (interaction) of mutations, and identify the mechanisms behind epistasis. Epistasis means that effects of mutations are context dependent (e.g. in the presence of other mutations), which severely restricts the trajectory of evolution. To tackle this, I created a complete enzyme evolution in the laboratory, and analyzed the effects of the mutations.

Directed evolution of PTE (phosphotriesterase) to esterase activity

The protein PTE was initially selected as a target of interest for performing evolutionary experiments in the laboratory, and a set of PTE variants carrying mutations was created. PTE catalyses phosphotriester hydrolysis with very high efficiency ($k_{cat}/K_M = 10^7$) and is also known to catalyze other promiscuous reactions such as lactonase and esterase ($k_{cat}/K_M \approx 10^{2\text{-}3}$). As a continuation of work initiated with Prof. Dan Tawfik at the Weizmann Institute of Science, I performed a directed evolution experiment of PTE to improve the promiscuous esterase activity. To make protein evolution efficient and continuous, I used a trick in the directed evolution process, co-expression of the chaperonins GroEL/ES. Since most mutations, and those that alter function in particular, are destabilizing, compensating and buffering the destabilizing effect of mutations was found to be crucial in the evolutionary process. GroEL/ES are known to assist folding of other proteins, and thereby were able to buffer the destabilizing effect of mutations. The over-expression of GroEL/ES was turned on when function altering mutations were desired. On the other hand, when a PTE variant was destabilized too much and the protein was expressed in inclusion bodies even with GroEL/ES over-expression, the over-expression of GroEL/ES was turned off, and stabilizing mutations were selected to increase the soluble protein expression level in the cell. Cycles of switching on and off GroEL/ES over-expression ensured that the PTEs were solubly expressed and evolvable, therefore the evolution proceeded smoothly and continuously.

With 18 cycles of directed evolution experiments, I successfully evolved PTE to have highly efficient esterase activity ($k_{cat}/K_M = 5\text{x}10^6$), and decreased the phosphotriesterase activity 10^4 fold (over–all a 10^8 fold specificity switch) with an accumulation of 18 mutations. It should be noted that this is one of the few experiments to completely switch the function of an enzyme in the laboratory, and provides very useful materials to tackle many questions in protein evolution: how can enzymes switch from one function to another, how does structure change throughout to the functional transition, how epistatic are mutations during evolution, and how is the evolutionary trajectory restricted?

The role of stability in functional evolution

To examine how stability restricted the evolutionary trajectory, I measured soluble expression level of variants in the evolution. The difference in enzymatic activity in crude lysates following expression with and without GroEL/ES over-expression (GroEL dependency) served as an indicator of the enzyme's stability. We performed the selection without GroEL/ES over-expression in the 3rd, 9th, 10th and 11th generations (highlighted in green in Figure 1), and selected four mutations (I274S, A80V, S111R and A204G) that reduced GroEL dependency, which indicates that these four mutations are stabilizing. SDS-PAGE analysis revealed that these mutations significantly improved soluble expression level in the cell without GroEL/ES over-expression. Moreover, I created single point

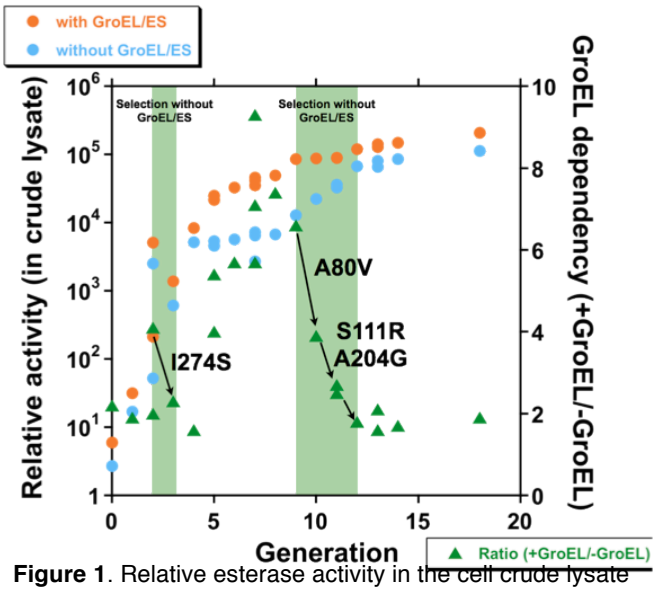

Figure 1. Relative esterase activity in the cell crude lysate of variants. GroEL/ES dependency is the ratio between the activity with and without GroEL.overexpression.

mutants of the four mutations on top of the evolved variant G19, and measured soluble expression. This revealed that the four mutations reduced soluble expression level significantly, and that the GroEL dependence decreased from 2 to 10 when all four mutations were accumulated on top of the G19. The biophysical experiments, however, failed to correlate heat stability and solubility of protein in the cell. Further studies to measure protein stability using a chemical denaturant are needed to to link the stability, solubility and evolvability of a protein. These results clearly demonstrate that stability plays a critical role in functional evolution, and causes epistasis since the effect of function and stabilizing mutations are context dependent, e.g. when a protein is not stable, function altering but destabilizing mutations cannot be fixed because they decrease the expression level of protein and therefore the fitness of protein. In the same way, stabilizing mutations are only positive when the protein is already unstable and poorly expressed in the cell, but are otherwise neutral because there is no direct advantage to excess stability. This function-stability tradeoff is expected to be a major constraint in evolution, and manipulating stability by buffering and stabilizing mutations is necessary to achieve large functional improvement by directed evolution, protein engineering and design.

Epistasis between functional mutations

I created all single point mutants of 12 functional mutations accumulated in the evolution on top of the wild type PTE sequence, and also single point revertant mutations (a mutation back to the original amino acid residue) on top of the evolved variant PTE-G19, and measured the context dependent effects of the mutations on catalytic activity. Interestingly, very strong epistasis was observed, e.g. H254R increased esterase activity 5-fold when it appeared in the evolution (first generation), however, when the mutation was reverted on top of G19 (R254H), it decreased the esterase activity 1000-fold. On the other hand, D233E increased the esterase activity 150-fold but only decrease it 3-fold when reverted from G19. Also some of mutations showed sign epistatsis, where the effect of the mutation is opposite in a different context. The mutations causing strong epistasis appeared to be located on one part of the structure and form an interaction network. Further detailed analysis with multiple combinatorial mutations and determination of their function and structure will be necessary to reveal the mechanism of the epistasis.

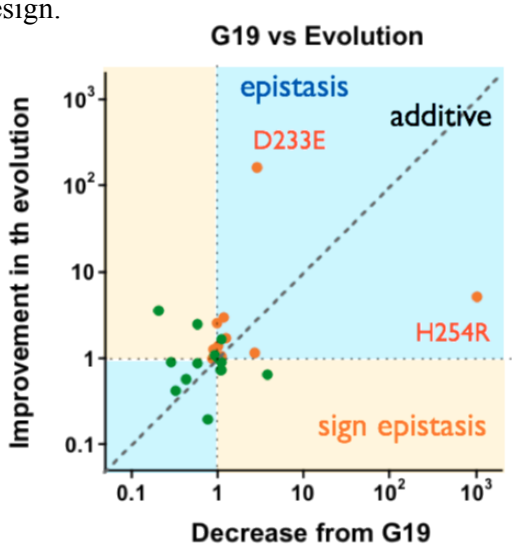

Figure 2. Comparison of context dependent activity difference of the effect of mutations.

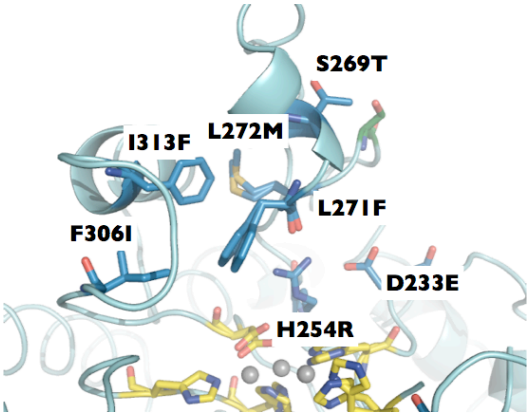

Figure 3. Position of mutations clustered in the one part of structure.

Summary

We successfully identified two types of epistasis in the enzyme evolution created in the laboratory; one is caused by a function-stability trade off, which strongly constrains the evolutionary pathway of the protein, and the other is due to physical interactions of amino acid residues. Although further deep analyses are required, this work has revealed mechanisms by which evolution is restricted and resulted in methods by which we can overcome these restrictions and engineer highly efficient proteins.