Proposal for an enzyme redesign method to improve production rates in *Aspergillus niger*

B.A. van den Berg1,3,4, M.J.T Reinders1,3,4, H.J. Pel2, J.A. Roubos3, D. de Ridder1,3,4

1 The Delft Bioinformatics Lab, Faculty of Electrical Engineering, Mathematics & Computer Science, Delft University of Technology, Delft, The Netherlands, 2 DSM Biotechnology Center, Delft, The Netherlands, 3 Netherlands Bioinformatics Centre, Nijmegen, The Netherlands, 4 Kluyver Centre for Genomics of Industrial Fermentation, Delft, The Netherlands

* b.a.vandenberg@tudelft.nl

Introduction

High yields are required for industrial production of enzymes. Previous work showed that in the microbial cell-factory *Aspergillus niger* a protein’s amino acid composition is predictive for high-level production1. To improve production rates of enzymes for which we did not observe high-level production, we propose a design method that increases resemblance to proteins for which high-level production was observed. Taking into account protein structure, our algorithm modifies the amino acid composition to better match that of structurally similar, but high-level produced proteins.

A. Structure prediction

Homology modeling software (ITASSER) is used to predict the tertiary structure based on the protein sequence, excluding the predicted signal peptide.

B. Mutation restrictions

All residues in the vicinity of the active side are fixed (colored sticks in structure B). At all other positions, only mutations are allowed that are also observed on the same position in homologous proteins and that are predicted to improve the thermostability of the protein.

Sequence position

Residue at this position

<table>
<thead>
<tr>
<th>Pos</th>
<th>T/A</th>
<th>Hom</th>
<th>ΔΔG</th>
</tr>
</thead>
<tbody>
<tr>
<td>91</td>
<td>R</td>
<td>KRNLARP</td>
<td>K</td>
</tr>
<tr>
<td>92</td>
<td>L</td>
<td>LH</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>V</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>W</td>
<td>WRYMVLL</td>
<td>YF</td>
</tr>
<tr>
<td>96</td>
<td>V</td>
<td>VFL1MY</td>
<td>AGST</td>
</tr>
<tr>
<td>97</td>
<td>H</td>
<td>FYIM</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>I</td>
<td>IFL</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Q</td>
<td>QLG</td>
<td></td>
</tr>
</tbody>
</table>

Allowed amino acids based on free energy calculation, only allowing for mutations that provide a decrease in free energy (negative ΔΔG).

C. Protein design

A protein for which high-level secretion was not observed is used as redesign target. The design method is based on three data sources: 1) the table from step B restricts what mutations are allowed at each position, 2) the amino acid contributions in Figure 1 define what mutations are desired, with a mutation from the most negative to the most positive contribution as the most favorable, and 3) the amino acid composition of 7 proteins that are structurally similar to the target, but for which high-level production was observed, puts boundaries on the amino acid composition. The last step ensures that the most favorable mutation (K -> N) is not selected too often, as this would result in a highly skewed amino acid composition.

References