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

van den Berg, Bastiaan; Reinders, Marcel; Pel, HJ; Roubos, J.A.; de Ridder, Dick

Publication date
2012

Document Version
Final published version

Citation (APA)
van den Berg, B., Reinders, M., Pel, HJ., Roubos, J. A., & de Ridder, D. (2012). *Proposal for an enzyme redesign method to improve production rates in Aspergillus niger*. 1.

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

B.A. van den Berg*[1,3,4], M.J.T Reinders[1,3,4], H.J. Pel[2], J.A. Roubos[2], D. de Ridder[1,3,4]

[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 production[1]. 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

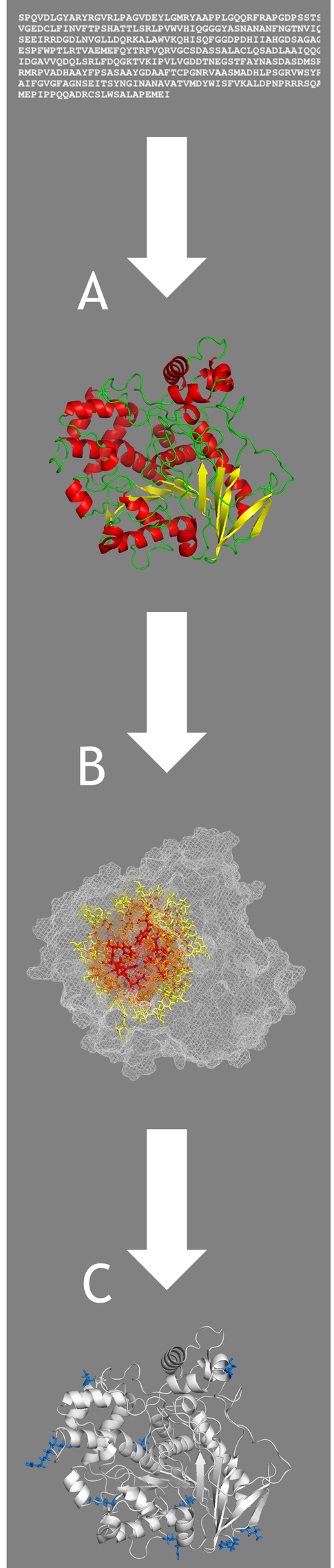
Residue at this position

Fixed residues: those that are predicted to be ligand binding or active and all residues that reside within 8 Å distance of those.

Allowed amino acids based on multiple sequence alignment with homologous proteins.

pos	r	A	hom	ΔΔG
.	.	.	.	.
91	R		KRNLAEP	K
92	L		LH	
93	C		C	
94	V		V	
95	W		WFYMVL	YF
96	V		VFLIMY	AGST
97	H		FYWH	
98	I	f	IFL	
99	Q	f	QLG	
.	.	.	.	.

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 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.

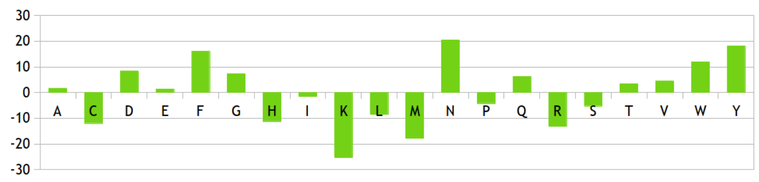


Figure 1 *- The bars denote the contribution of the different amino acids to successful high-level production, as obtained in previous work. For example, asparagine (N) and tyrosine (Y) have a positive influence on high-level production, whereas lysine (K) has a negative influence.*

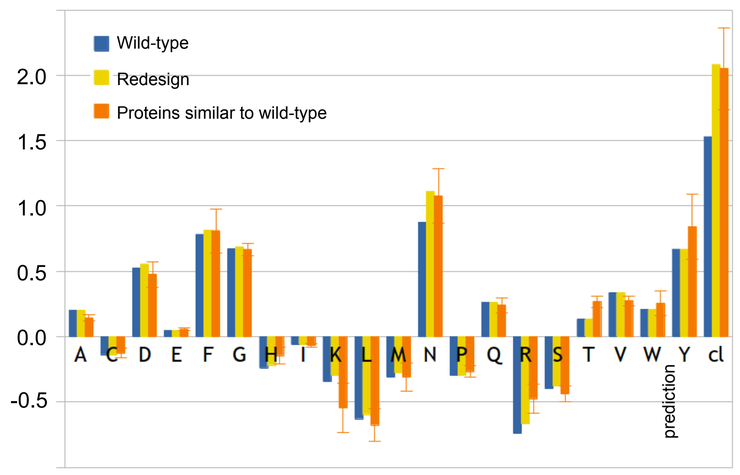


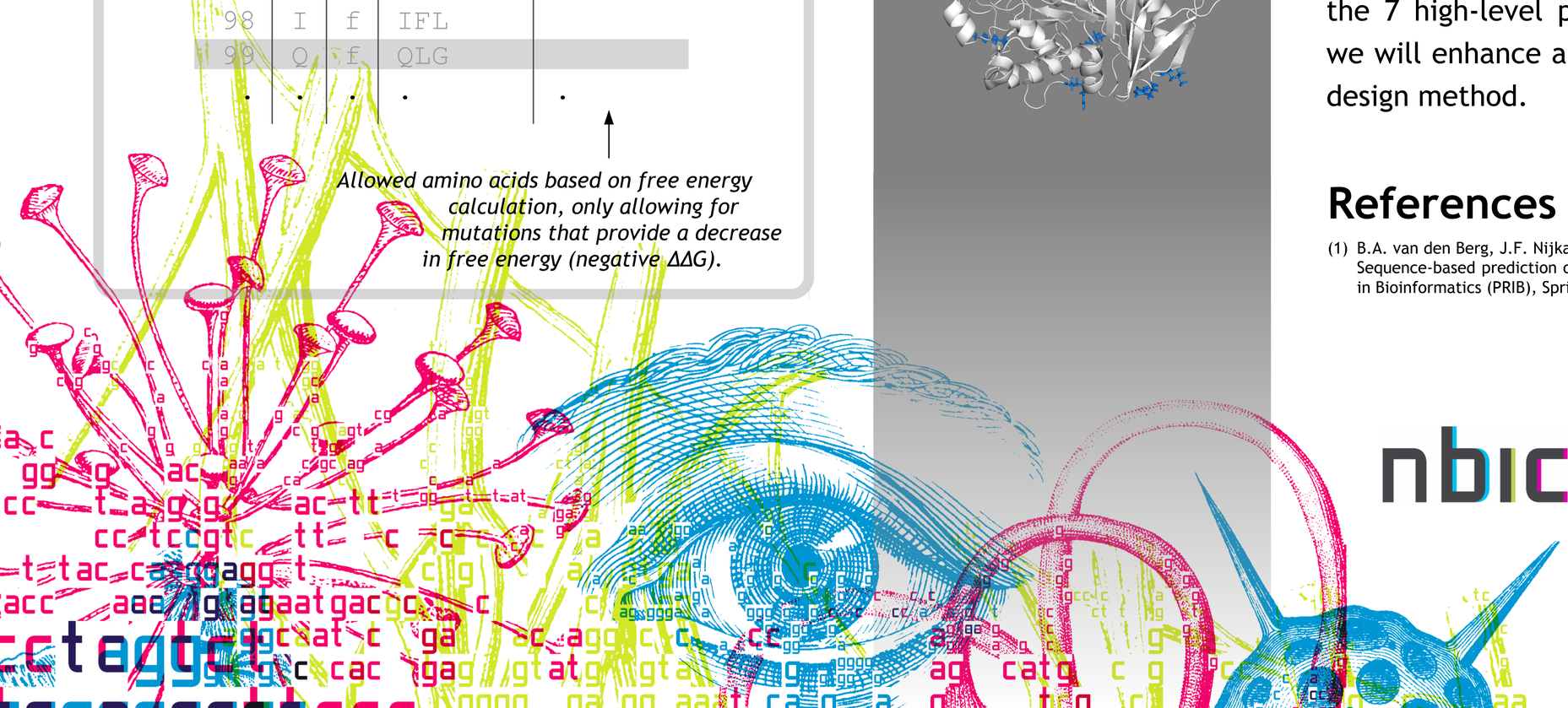
Figure 2 *- The amino acid composition weighted by the amino acid contributions from Figure 1. The last three bars are the prediction scores for high level production. Blue: original target protein. Orange: redesigned target protein. Yellow: average (+- standard deviation) of 7 proteins similar to the target protein for which high-level production was observed.*

Conclusion

Initial test runs of the proposed algorithm indicate that a limited number of mutations (~10) are needed to obtain a prediction score for high-level secretion that is similar to the 7 high-level produced proteins. In the near future, we will enhance and experimentally validate our rational design method.

References

(1) B.A. van den Berg, J.F. Nijkamp, M.J.T. Reinders, L. Wu, H.J. Pel, J.A. Roubos, and D. de Ridder. Sequence-based prediction of protein secretion success in *Aspergillus niger*. Pattern Recognition in Bioinformatics (PRIB), Springer Lecture Notes in Bioinformatics vol. 9282, 3-14, 2010.



nbic netherlands bioinformatics centre

TU Delft

DSM BRIGHT SCIENCE. BRIGHTER LIVING.

Kluyver CENTRE | Kluyver Centre for Genomics of Industrial Fermentation