COMPUTATIONAL OPTIMIZATION OF GLYCEROL DEHYDROGENASE FOR INDUSTRIAL APPLICATIONS

Ricardo V. L. Zappala[1*], Patrick N. Squizato [1], Carlos A. X. Gonçalves [1] & Ana C. A. Collaço [1]

[1] *Biotechnology department, SENAI Innovation Institute for Biosynthetics and Fibers, Rio de Janeiro, Brazil.*
** rladewig@cetiqt.senai.br*

ABSTRACT

The industrial applications of glycerol dehydrogenase (glda) can range from biorrefinaries and circular economy to production enhancement and lowering costs. When searching for protein alternatives, one should always consider a computational step for optimization, that saves thousands of dollars during experimental trials. The present work shows that glda may be enhanced using database references and computational tools, regarding both its sequence and structure. We investigated glda sequence in enzyme databases, evaluated possible site-directed mutants and tested them using molecular docking. A protein ranking was formed to select the best candidates to be carried over. The rational design of such protein represents a cheaper way to develop new enzymes for the industry, taking the best in silico candidates to validation essays and accelerating protein development itself.

Keywords: enzymes. bioinformatics. Molecular docking.

1 INTRODUCTION

Protein engineering can generate several candidates for industrial applications, combining or selecting between rational design and directed evolution (Maeda *et* al., 2008). Previous reports have shown that glda could be rationally designed for enzyme improvement considering site-directed mutagenesis in *E. coli* (Zhang *et* al., 2010).

The present work is the first step towards producing an optimal enzyme to convert glycerol into dihydroxyacetone (dha). Using computational tools as the initial workflow and filter, a list of predicted enzyme modifications was created ranking several candidates. These proteins will be carried over to experimental trials and bioprocess optimization in later steps.

2 MATERIAL & METHODS

To stablish a solid optimization pipeline, several approaches were taken considering both sequence and the protein structure. *Escherichia coli* enzyme was chosen for its quick response and versatility in later bench experiments. Moreover, once the pipeline was valid for a model organism as *E. coli*, several others could be improved with a shorter learning process as it is necessary for innovation.

Firstly, glda redocking was conducted using autodock vina and Autodock4Zn forcefield (Santos-Martins *et* al., 2014). The main structure for the enzyme tests was based on PDB ID:5ZXL and NAD+ from PDB ID: 1JQ5. Chimera X (Meng EC *et* a., 2023) was used for protein-ligand complex visualization and handling. This method would serve as a reference for site-directed mutated proteins and their molecular docking tests.

Secondly, the promissing mutants for glda were prospected following the best reported results for ligand affinity and stability. BRENDA database (Chang *et* al., 2021) was consulted, comparing several organisms and compiling total information into a dataframe for proper visualization and decision-making. For mutant suggestions, the main webserver used was HotSpot Wizard (Sumbalova *et* al., 2018) and Consurf (Ashkenazy *et* a., 2016) for residue conservation analysis.

After the viable mutants were listed, their structures were prepared for molecular docking considering both glycerol and dha as ligands. The final ranking was constructed using both autodock vina outputs and previous sequence analysis. The best theoretical reactions and structures would then have their corresponding genes synthetized for *in vitro* experiments.

RESULTS & DISCUSSION

Using Autodock Vina module for metalloproteins (Zn), the native enzyme system was prepared to generate basic estimated energies for both glycerol and dha ligands. Later, these redocking values would be used as a reference for new proteins simulations.

Investigating BRENDA database and complementary information, we found that different mutation strategies were conducted to enhance glda in other species. A total of 29 species were compared considering reaction parameters for this enzyme, including optimal temperature, Km and pH. We observed that extremophile species express promising glda enzymes, although not always in the best industrial conditions. Moreover, the lack of information regarding dha substract would increase the risk of manipulating species other than *E. coli*.

To design *E. coli* glda proteins, we considered inputs from previous reports of the original species and extremophile ones. We also gathered insights from prediction servers as hotspot wizard and PROSS, avoiding conserved residues and forming 9 mutant sequences total.

As shown in **Figure 1**, several complexes were visualized in Chimera X and both their RMSD and predicted ligand energies were analyzed to rank each protein. After this, the best proteins were selected to be carried over in experimental validation. Since we wanted to optimize the glycerol convertion and decrease the opposite reaction, a gli/dha ratio indicated that mutants 2, 7 and 8 would be the most promising ones.

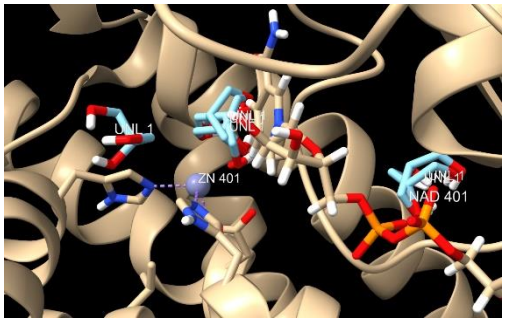

Figure 1. Example of vina autodock report with complex visualization in Chimera X for glycerol (UNL). Zinc and NAD+ are indicated by labels in the catalytic site region.

CONCLUSION

After computational optimization of glda enzyme, we conclude that three candidates should be prioritized in experimental tests. Mutant 2 for its predicted decrease in glycerol formation, which is an important countering variation, mutant 7 for its predicted increase in glycerol formation and mutant 8 considering the best glycerol/dha ratio when comparing docking energies.

REFERENCES

[1] Maeda T, Sanchez-Torres V, Wood TK 2008. Appl Microbiol Biotechnol 79:77–86
[2] H. ZHANG, C. B. CHING, R. JIANG 2010. Appl. Microbiol. Biotechnol. 88, 117–124
[3] Santos martins 2014. J. Chem. Inf. Model. 54, 2371−2379.
[4] Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE 2023. Protein Sci. 32(11):e4792.
[5] Chang A., Jeske L., Ulbrich S., Hofmann J., Koblitz J., Schomburg I., Neumann-Schaal M., Jahn D., Schomburg D. 2021. Nucleic Acids Res.*, 49:D498-D508*.
[6] Sumbalova, L., Stourac, J., Martinek, T., Bednar, D., Damborsky, J. 2018. Nucleic Acids Research 46 (W1): W356-W362
[7] aAshkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. Nucleic Acids Res.44(W1):W344-50

ACKNOWLEDGEMENTS

The work described herein was supported by SENAI CETIQT and SENAI Innovation Institute collaborators. All authors contributed to this work from its inception to its consolidation.