HOMOLOGY MODELLING AND DOCKING ANALYSIS OF L-LACTATE DEHYDROGENASE FROM STREPTOCOCCUS THERMOPILUS

Vladimir R. Vukić*, Spasenija D. Milanović, Mirela D. Iličić, Katarina G. Kanurić and Dajana V. Vukić

University of Novi Sad, Faculty of Technology, Bulevar Cara Lazara 1, 21000 Novi Sad, Serbia

The aim of this research was to create a three-dimensional model of L-lactate dehydrogenase from the main yoghurt starter culture – Streptococcus thermopilus, to analyse its structural features and investigate substrate binding in the active site. NCBI BlastP was used against the Protein Data Bank database in order to identify the template for construction of homology models. Multiple sequence alignment was performed using the program MUSCLE within the UGENE 1.11.3 program. Homology models were constructed using the program Modeller v. 9.17. The obtained 3D model was verified by Ramachandran plots. Molecular docking simulations were performed using the program Surflex-Dock. The highest sequence similarity was observed with L-lactate dehydrogenase from Lactobacillus casei subsp. casei, with 69% identity. Therefore, its structure (PDB ID: 2ZQY:A) was selected as a modelling template for homology modelling. Active residues are by sequence similarity predicted: S. thermophilus – HIS181 and S. aureus – HIS179. Binding energy of pyruvate to L-lactate dehydrogenase of S. thermopilus was -7.874 kcal/mol. Pyruvate in L-lactate dehydrogenase of S. thermopilus makes H bonds with catalytic HIS181 (1.9 Å), as well as with THR235 (3.6 Å). Although our results indicate similar position of substrates between L-lactate dehydrogenase of S. thermopilus and S. aureus, differences in substrate distances and binding energy values could influence the reaction rate. Based on these results, the L-lactate dehydrogenase model proposed here could be used as a guide for further research, such as transition states of the reaction through molecular dynamics.

KEY WORDS: L-lactate dehydrogenase, Streptococcus thermopilus, homology modelling, molecular docking

INTRODUCTION

L-lactate dehydrogenase (EC: 1.1.1.27) belongs to the family of 2-hydroxy acid oxidoreductases and catalyses the conversion of pyruvic acid to lactate and back, as it converts NADH to NAD⁺ and back. Anaerobe microorganisms use lactate dehydrogenase to oxidise NADH to NAD, which is an essential reaction in fermentation processes.

* Corresponding author: Vladimir Vukić, University of Novi Sad, Faculty of Technology, Bulevar Cara Lazara 1, Novi Sad, Serbia, vukicv@tf.uns.ac.rs
Structural analysis of lactate dehydrogenase from microorganisms include several crystal structures: Staphylococcus aureus (unpublished work), Bacillus stearothermophilus (1), Bifidobacterium longum (2), Lactobacillus casei (3), Plasmodium falciparum (4), etc. The diversity of amino acid sequences between each of these proteins is relatively low. Also, high level of structural similarity is preserved. Regarding milk fermentation, there are crystal structures of lactate dehydrogenases from some strains used as starter cultures - Lactobacillus casei subsp. casei (3) and Bifidobacterium longum subsp. longum (2). However, crystal structure of the L-lactate dehydrogenase from yoghurt starter culture Streptococcus thermophilus is not known. Molecular modelling methods have been successfully used in research of enzyme structure, reaction mechanisms, and engineering (5). Modelling methods include prediction of protein stability, enzyme specificity and selectivity, mechanism, solubility, and interaction with interfaces. Molecular modelling methods were used to investigate the relationship between the structure of enzymes and their biochemical properties (6). Numerous approaches have been tested to stabilise proteins and improve their activity by using modelling strategies such as optimisation of electrostatic interactions (7) improvement of side chain packing (8), or balancing the ratio between hydrophilic and hydrophobic residues (9). The molecular modelling methods have been successfully applied to engineering a lipase into an aldolase (10) or into enzymes that catalyse Baeyer–Villiger oxidation with hydrogen peroxide (11). Therefore, the aim of this research was to create a three-dimensional model of L-lactate dehydrogenase from the main yoghurt starter culture - Streptococcus thermophilus (strain CNRZ 1066), to analyse its structural features and investigate substrate binding in the active sites.

MATERIALS AND METHODS

NCBI BlastP [12] was used against the Protein Data Bank (PDB) database [13] in order to identify template L-lactate dehydrogenase structures with the highest sequence similarity to the L-lactate dehydrogenase sequences from Streptococcus thermophilus. Because the present study focuses on the initial steps in substrate binding to L-lactate dehydrogenase, we selected template crystal structure without substrates in the active site. Three-dimensional coordinates of L-lactate dehydrogenase used as template structure was retrieved from PDB data base (PDB ID: 2ZQY:A). The molecular structure of lactate and NADH were extracted from crystal structure of L-lactate dehydrogenase from Staphylococcus aureus subsp. aureus (PDB ID: 3H3J:A). Energy minimization was performed using a Trypos force field. Protein sequences for alignment and construction of homology model were retrieved from the National Center for Biotechnology Information (NCBI) (Lactobacillus casei subsp. casei - GenBank: BAA02133.1; Streptococcus thermophilus - GenBank: AAV62824.1; Staphylococcus aureus subsp. aureus - GenBank: AAW38779.1). Multiple sequence alignment was performed using the program MUSCULE within the UGENE 1.11.3 program (14, 15). Highly conserved regions and potential active site residue were identified by sequence alignment.

Homology models were constructed using the program Modeller v. 9.17 (16). The best model was selected by considering the smallest value of the normalised discrete optimized molecule energy (DOPE) (17). The obtained 3D model was verified by Rama-
chandran plots and using the Structural Analysis and Verification Server (18). The modelled protein structure was subjected to energy minimization using an AMBER7 force field and AMBER-derived atomic charges. The Powell method, distance dependent dielectric constant and convergence gradient method with a convergence criterion of 0.005 kcal/mol were applied. The modelled protein structure was prepared for molecular docking simulation using the following parameters: polar hydrogen atoms were added, protonation types and termini treatment was enabled. Docking simulations were performed using the program Surflex-Dock with flexible H atoms (19). The docking results were visualised using the program PyMol (20).

RESULTS AND DISCUSSION

BlastP analysis of L-lactate dehydrogenase from S. thermophilus revealed a high degree of sequence similarity with several X-ray crystal structures retrieved from the PDB. The most important PDB structures are presented in Table 1. The associated E-values confirm a correlation between query and reference sequences. The highest sequence similarity was observed with L-lactate dehydrogenase from Lactobacillus casei subsp. casei, with 69% identity. Therefore, its structure (PDB ID: 2ZQY:A) was selected as a modelling template for homology modelling. The structure of L-lactate dehydrogenase from Staphylococcus aureus subsp. aureus was experimentally determined with both substrates (NADH and pyruvate) in the active site (PDB ID: 3H3J:A). Therefore, it reveals the early steps of substrate recognition. It has high sequence similarity with the template sequence (58%), which makes it appropriate for evaluations of the docking results.

Table 1. BlastP results of query L-lactate dehydrogenase sequence against the PDB

<table>
<thead>
<tr>
<th>Query sequence</th>
<th>PDB sequence</th>
<th>Query cover</th>
<th>E value</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. thermophilus</td>
<td>2ZQY:A (Lactobacillus casei subsp. casei)</td>
<td>95%</td>
<td>3e^-159</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>3H3J:A (Staphylococcus aureus subsp. aureus)</td>
<td>93%</td>
<td>1e^-119</td>
<td>58%</td>
</tr>
</tbody>
</table>

* E value - probability of the alignment occurring by chance

Active residue of L-lactate dehydrogenase of L. casei has been determined as HIS181 (21). The sequence alignment revealed highly conserved HIS residue which acts as an active residue (Fig. 1). The active residues are by sequence similarity predicted: S. thermophilus – HIS181 and S. aureus – HIS179. Furthermore, there are highly conservative regions around the active residue, which is crucial for highly accurate homology modelling of the enzyme.
Figure 1. Sequence alignment of active site residues from the selected L-lactate dehydrogenases

Homology models of L-lactate dehydrogenase from *S. thermophilus* was constructed using Modeller program and the best model was chosen using the DOPE value (results not shown). Ramachandran plots were constructed for the modelled enzyme (Fig. 2). Based on the Ramachandran plot analysis, most residues of the modelled enzyme are in the favoured region, and only a few residues are in unfavourable conformations (Fig. 2a). A detailed analysis of residues in the active site revealed that all residues involved in substrate bindings are in the favoured region (Fig. 2b). These results revealed that the enzyme residues have favourable bond angles, which is extremely important for molecular docking simulations, and suggests that the model is suitable for analysing of substrates binding.

Figure 2. Ramachandran plot validation of the *S. thermophilus* L-lactate dehydrogenase homology model. Plot of phi/psi angles (-150→150) vs. violation (gradient, blue→red) a) homology model of *S. thermophilus* L-lactate dehydrogenase; b) Active site residues of *S. thermophilus* L-lactate dehydrogenase

The modelled enzyme was further validated by structural superposition with its corresponding template sequence (Fig. 3). The model showed a high degree of three-dimensional similarity with its template structure, with the Cα Root Mean Square Deviation (RMSD) of 0.78Å.
Using our molecular docking protocol, pyruvate was successfully re-docked with the crystal structure of *Staphylococcus aureus* subsp. *aureus* L-lactate dehydrogenase (RMSD of 0.249 Å), suggesting that the methods used in the present study are appropriate. Molecular docking simulation revealed position of substrates in the active site of L-lactate dehydrogenase (Fig. 4a). The binding energy of pyruvate to L-lactate dehydrogenase of *S. thermopilus* was -7.874 kcal/mol, while binding energy to L-lactate dehydrogenase of *S. aureus* was -7.325 kcal/mol. These results indicate stronger binding of pyruvate to the modelled enzyme than to *S. aureus*. The pyruvate in L-lactate dehydrogenase of *S. thermopilus* forms H bonds with catalytic HIS181 (1.9 Å), as well as with THR235 (3.6 Å). Threonine residue is also involved in pyruvate binding of some other L-lactate dehydrogenases with high sequence similarity with the enzyme of *S. thermopilus*, including L-lactate dehydrogenase from *S. aureus* subsp. *aureus* (Fig. 4b). A comparison of our molecular simulation results with the crystal structure of L-lactate dehydrogenase from *S. aureus* subsp. *aureus* revealed similar substrate positions. Nevertheless, the pyruvate in the enzyme from *S. thermopilus* is closer to the catalytic residue HIS181 than to the substrate binding residue THR235, while NADH is at a similar distance from the catalytic residue. Although our results indicate similar position of the substrates, the differences in substrate distances and binding energy values could influence the reaction rate. Although predicted docking energies are not reliable for modelling enzyme efficiency, they do correlate with protein-ligand complex stability and overall reaction rates. Based on these results, the L-lactate dehydrogenase model proposed here could be used as a guide for further research. Detailed reaction mechanism could be investigated through molecular dynamics simulations. Quantum mechanical calculations could reveal reactions highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) which could identify transition states of the reaction. These results could lead to rational design and desired modifications of the enzyme that could be used for production of lactic acid.
Figure 4. Substrate binding (pyruvate and NADH) to L-lactate dehydrogenase: a) Molecular docking simulation of *S. thermophilus* (homology model); and b) crystal structure of *Staphylococcus aureus* subsp. *aureus* (PDB ID: 3H3J:A)

**CONCLUSION**

The BlastP analysis of L-lactate dehydrogenase from *S. thermophilus* revealed the highest degree of sequence similarity with X-ray crystal structure of L-lactate dehydrogenase from *Lactobacillus casei* subsp. *casei* with 69% identity. Active residues are by sequence similarity predicted: *S. thermophilus* – HIS181 and *S. aureus* – HIS179. Homology models of L-lactate dehydrogenase from *S. thermophilus* was constructed and validated. The modelled structure showed a high degree of three-dimensional similarity with its template structure, with the Ca RMSD of 0.78 Å. Molecular docking simulation revealed that pyruvate forms H bonds with catalytic HIS181 (1.9 Å), as well as with THR235 (3.6 Å). Although our results indicate similar position of substrates of L-lactate dehydrogenase from *S. thermophilus* and *S. aureus* subsp. *aureus*, the differences in substrate distances could influence the reaction rate. Based on these results, the L-lactate dehydrogenase model proposed here could be used for molecular dynamics simulations and quantum mechanical calculations that could lead to rational design of the enzyme and improvement of lactic acid production.

**Acknowledgement**

The authors thank the Ministry of Education, Science and Technological Development of the Republic of Serbia for financial support of the present study (Project No. III 46009).
REFERENCES


3. Arai, K; Ishimitsu, T; Fushinobu, S; Uchikoba, H; Matsuzawa, H; Taguchi, H. Active and inactive state structures of unliganded Lactobacillus casei allosteric L-lactate dehydrogenase. Proteins 2010, 78, 681-694.


Циљ овог истраживања био је конструкција тродимензионалног модела L-лактат дехидрогеназе S. thermopilus, као и анализе структуралних особина и везивање субстрата у активном месту. NCBI BlastP претрага PDB базе података је коришћена за проналажење шаблон структуре за конструкцију хомологих модела. Вишеструко поравнање секвенци је урађено помоћу програма MUSCLE у оквиру UGENE 1.11.3 програма. Хомологи модели су конструисани помоћу програма Modeller v. 9.17. Добијени модели су верификовани кроз Рамачандранове плотове. Симулације молекуларног докинга је урађена применом Surflex-Dock програма. Највећа сличност секвенци је пронађена са L-лактат дехидрогеназом из Lactobacillus casei подврста casei, са сличношћу од 69%, услед чега је одабрана као шаблон за конструкцију хомологих модела (PDB ID: 2ZQY:A). Активне резидуе се предвиђене путем сличности секвенци: S. thermopilus – хистидин 181 и S. aureus – хистидин 179. Енергија везивања пирувата за активно место L-лактат дехидрогеназе S. thermopilus-a износи -7,874 kcal/mol. Пируват гради водоничне везе са каталитичком резидуом хистидина (1,9 Å), као и са треонином 235 (3,6 Å). Иако наши резултати указују на сличне позиције супстрата - лактат дехидрогеназа S. Thermopilus-a и S. aureus-a, разлике у удаљеностима супстрата од резидуа, ако и енергија везивања, указују на могуће разлике у брзини реакције. Резултати добијени у овом раду могу бити коришћени у даљим истраживањима, као што је идентификација прелазних стања реакције применом молекулярне динамике. 

Кључне речи: L-лактат дехидрогеназа, Streptococcus thermopilus, хомолого моделирање, молекулярни докинг