Virtual screening for the discovery of novel urease inhibitors of rumen bacterial urease

Shengguo Zhao[1*], Norovsambuu Togtokhbayar[2], Baatar Narantuya[3]

[1] Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing, China; [2] School of Animal Science and Biotechnology, Mongolian University of Life Science, Mongolia, Ulaanbaatar, Mongolia; [3] Research Institute of Animal Husbandry, Ulaanbaatar, Mongolia

[*] Corresponding author: Shengguo Zhao, zhaoshengguo@caas.cn.

Abstract: Inhibition of urea hydrolysis by rumen bacterial urease can improve the efficiency of dietary nitrogen utilization in animal production and decrease nitrogen pollution into the environment. The three dimensional structure of urease from the ruminal bacterium Ruminococcus albus 8 (RaUrease) was resolved by comparison with the crystal structure of urease from Helicobacter pylori (HpUrease, PDB ID 1E9Z) using Modeller. A reliable structure of RaUrease was generated using molecular dynamics simulation for 20 ns with the AMBER force field using the Gromacs system, and the stereochemical quality of the protein model was evaluated. On the basis of binding energies (<-9.8 kcal mol-1), 12 commercially available agents were selected as potential urease inhibitors. Characterization of the RaUrease-binding site for the compound with the highest predicted inhibitory activity indicated that van der Waals forces were more important than hydrogen bonds for rumen bacterial urease binding. The results of the present study may help design novel urease inhibitors for ruminants and alleviate nitrogen pollution derived from livestock

production.

Keywords: *Ruminococcus albus* 8; urease inhibitors; virtual screening; molecular docking; molecular dynamics

1. Introduction

Urease (urea amidohydrolas, EC 3.5.1.5) is a thiol-rich and nickle-dependent metalloenzyme than can catalyze the hydrolysi of urea, thereby producing ammonia and carbamate. The nickel ions (Ni^{2+}) and the sulfhydryl group, especially the multiple cysteinyl residues in the active site of the enzyme, are essential for the catalytic activity of all urease. Structural studies of HPU have revealed a dinuclear Ni active site with a carbamylated lysine residue that bridges the deeply buried metal atoms. Around the active site, the hydrophilicity of amino acids is characterized by highly flexible flap that undergoes an induced fit. It has been revealed that the urease catalytic activity strongly depends on its multiple cysteinyl residues bearing sulfhydryl groups, especially those that is located on the mobile flap closing the active site.

Urea has been used as a supplement in ruminant rations for more than 100 years to increase ruminal degradable nitrogen and reduce feed cost [1]. However, the practice of feeding urea to ruminants should be approached with caution. Because, high rate of urea hydrolysis induced by bacterial urease in rumen can exceed the capacity of ammonia assimilation, which will lead to a low efficiency of nitrogen utilization, animal blood ammonia toxicity, and nitrogen pollution in the environment [2,3]. Thus, varieties of urease inhibitors, acting as an adjunct to dietary urea, have been investigated in animal husbandry during past decades, such as acetohydroxamic acid (AHA), phenyl phosphorodiamidate (PPDA), and N-(n-butyl-)thiophosphoric triamide (NBPT).

Unfortunately, One strategy to reduce these negative effects is the use of urease inhibitors as an adjunct to dietary urea. Based on their inhibition efficiency and their effect on animal performance and productivity, urease inhibitors used in ruminant feeding practice include acetohydroxamic acid (AHA), phenyl phosphorodiamidate (PPDA), and N-(n-butyl-)

thiophosphoric triamide (NBPT). These inhibitors provide short-term control, since their efficacy decreases with time because of microbial adaptation [4-7]. Hence, the identification of novel and safe urease inhibitors with long-term activity is important in dairy production.

The discovery and validation of novel inhibitors for ruminants requires long-term livestock experiments, which are costly and time-consuming [8]. It is therefore important to narrow the scope of candidates before they are evaluated as animal feed. Virtual screening is a complementary approach to discover novel and potential hits from databases containing various commercially available compounds by docking them into the active site of the target protein rapidly and reliably [9]. This approach can dramatically reduce the number of potentially available compounds before they are tested in practice [10]. For this method, a model rumen bacterial urease with known three dimensional (3D) structure would be desirable as a target protein.

However, the isolation and identification of rumen ureolytic bacteria are difficult, and our knowledge of the structural and functional aspects of ureases from various ruminal bacteria is limited. Among the rumen bacterial community, *Ruminococcus albus* 8 is an important bacterium that is involved in the first step of cellulose decomposition and microbial protein synthesis. Bioinformatics analysis of the genomic sequence of this bacterium indicates that it can synthesize urease [11]; however, the urease from this bacterium has not been characterized, and its crystal structure is not available in the protein databank. Data from previous studies indicate that ureases from different bacteria exhibit high amino acid sequence homology, and the active sites of known bacterial ureases are located in the α subunits [12]. Comparative homology modeling provides a useful 3D model of proteins related to at least one protein with a known structure in the absence of

an experimentally determined structure [13]. Homology modeling can therefore be applied to generate a 3D structure of rumen ureolytic bacterial urease.

In the present study, homology modeling was used in combination with molecular dynamics simulations to obtain the 3D structure of urease from *R. albus* 8 (RaUrease). Molecular docking was performed to predict the binding site of RaUrease with its known inhibitor as a positive control. Finally, novel and potential inhibitors of rumen bacterial urease were selected using virtual screening. The results are expected to narrow the candidates for long-term animal experimental verification in the further.

2. Results and Discussion

2.1 Sequence alignments and molecular modeling

RaUrease contains 570 amino acids (Figure 1). Compared with the template sequence of HpUrease (568 amino acids), the sequence similarity between the two proteins was 69%. High accuracy comparative models are defined as more than 50% of sequence identity to their templates, which implies that model proteins have an approximately 1 Å root mean square error for the main chain atoms. Most errors are based on side chain packing and small shifts or distortions in the core main chain regions [14]. Thus, the 69% sequence similarity in the current study represents high homology to construct a reliable model. The predicted 3D model of RaUrease was obtained using Modeller 9.16, as shown in Figure 2 (a).

2.2 Optimization and validation of the homology model

In the molecular dynamics simulation study, conformational changes in the model structure of RaUrease were analyzed for a simulation period of 20 ns, and the predicted result is shown in Figure 2 (b). Ramachandran plot calculations were performed to evaluate the stereochemical quality of the

final refined model using the initial model structure of RaUrease as the reference structure. The detailed residue to residue stereochemical quality of the protein structure was assessed using the Molprobity program, and the results are shown in Figure 3 and Table 1.

Before the optimization, the percentages of residues located in the favored, allowed, and disallowed regions were 89.4% (508/568), 97.5% (508/568), and 2.46% (508/568), respectively. After the refinement of the model protein, the percentages of residues in the favored, allowed, and disallowed regions were 88.2% (500/567), 98.6% (500/567), and 1.44% (500/567), respectively. There were 14 outliers (VAL17, LYS49, MET55, CYS66, ALA100, MET192, ILE287, THR301, ASN302, ASP326, ASN328, MET367, VaL427, and ALA433) in the initial 3D structure of RaUrease. Eight outliers (ILE2, THR168, ASN302, PRO303, ASN465, CYS520, GLU560, and PHE668) were still present after 20 ns of molecular dynamics simulation. The ERRAT scores changed from 82.740 to 90.556. The percentage of residues with average three dimensional-one dimensional (3D-1D) scores ranged from 86.14% to 89.28%, according to the results of Verify-3D. Generally, after the optimization, quality factors increased and error values decreased by satisfying the special restraints.

Ramachandran plot calculations were performed on the structure of the template urease from *H. pylori*, which was used to evaluate the reliability of the final refined model structure of RaUrease (Table 1). The percentages of residues for *H. pylori* urease located in the favored, allowed, and disallowed regions were 75.5% (426/564), 92.7% (426/564), and 7.27% (426/564), respectively. The ERRAT score was 85.536, and 86.62% of residues in HpUrease had average 3D-1D scores. The results showed that the geometric quality of the backbone conformation, residue interactions and contacts, and energy profiles of the final model urease were similar to those of the template protein,

which suggests that a reasonable homology model of RaUrease was established. The verified model can be used for further investigations, such as protein-inhibitor interactions and virtual screening.

2.3 Identification of the binding region in the model urease by molecular docking

The ideal inhibitors are compounds that can fit well within the binding pocket. AHA is a simple and stable organic compound that has been used as a urease inhibitor in ruminant production in China since 1998 (No. 27 Letter on husbandry authorized by the Ministry of Agriculture of China). This representative inhibitor was docked to RaUrease using AutoDock Vina for the identification of the binding region in this model urease.

The RaUrease residues involved in the AHA-binding site are shown in Figure 4. AHA was surrounded by ALA390, LEU391, ASP394, SER395, GLN396, ASN398, ASP399, ASN400, PHE401, and ARG402. There were five hydrogen bonds formed between AHA and RaUrease. Among them, ASN398 formed two hydrogen bonds, and LEU391, PHE401, and ASP394 formed one each. The results imply that hydrogen bonds play an important role in AHA binding. The docking score for AHA was -4.5 kcal mol^{-1}, which was used as the reference score for further virtual screening. The AHA-binding region on RaUrease was used as one of the active sites for further virtual screening. To identify additional potential inhibitors, the grid size for AutoDock Vina in the next molecular docking was set to cover the entire model enzyme structure.

2.4 Selection of potential urease inhibitors and docking of the highest affinity inhibitor to RaUrease

The high-throughput virtual screening yielded 22 candidate compounds with free binding energies of <-9.8 kcal mol^{-1}. The 12 top hits are listed in Table 2. In comparison with the AHA score (-4.5 kcal mol^{-1}), these screened candidates with better docking scores should have stronger interactions and better inhibitory efficiency. Some of the compounds among the 12 candidates were

enantiomers, including the most efficient inhibitors, ZINC67911797 (-10.6 kcal mol^{-1}) and ZINC67911804 (-10.1 kcal mol^{-1}). To compare the two compounds, the binding patterns of ZINC67911797 and ZINC67911804 to the model RaUrease were determined and are shown in Figure 5.

The amino acid residues surrounding ZINC67911797 and ZINC67911804 were conserved. The two compounds contained the same 26 residues (HIS137, HIS139, ALA170, THR171, THR172, CYS173, LYS220, HIS222, GLU223, HIS249, THR252, HIS275, GLY278, ALA284, PRO285, LEU319, CYS322, HIS323, ARG339, ARG341, ASP363, and TYR544); however, the most effective inhibitor, ZINC67911797, had one more amino acid residue (ILE221) than ZINC67911804. There were five hydrogen bonds formed between ZINC67911797 and RaUrease, and seven hydrogen bounds formed between ZINC67911804 and RaUrease. These results suggest that van der Waals forces play a more important role than hydrogen bonds in the urease-binding site. Compared with the binding model of AHA, differences in the binding site and major intermolecular forces were found to affect the efficacy of the inhibitors.

3. Experimental Section

3.1 Homology modeling

The amino acid sequence of the target protein, RaUrease, was retrieved from the NCBI database (GenBank: EGC01992.1) in FASTA format. According to the results of the BLASTp algorithm, the template protein selected was the *Helicobacter pylori* urease (HpUrease) (PDB ID: 1E9Z). The sequence alignment was performed using Clustal X 2.0 software [15], and the 3D homology model structure of RaUrease was generated using Modeller 9.16 [16].

3.2 Molecular dynamics simulation

Molecular dynamics simulation was performed using Gromacs 5.1 with AMBER-03 all-atom force field [17]. The temperature was maintained at 300 K. The protein was solvated using a box of simple point charge (SPC) water molecules extending at least 10 Å away from the boundary of any protein atom. An integration step of 2 fs was used. Non-bonded interactions were calculated using a cutoff of 8 Å. Long-range electrostatic interactions were calculated by Particle–Mesh Ewald summation with grid spacing of 1.2 Å and cubic interpolation. After 1000 steps of steepest descent energy minimization, the solvent and ions were equilibrated by 0.5 ns molecular dynamics simulation with the protein heavy atoms subjected to harmonic constraints under a force constant of $k = 1000\ \mathrm{kcal \cdot mol^{-1} \cdot nm^{-2}}$. Finally, the production run was performed for 20 ns, and the coordinates of all the atoms were stored every 100 picoseconds for further analysis.

The model quality was assessed based on the geometric quality of the backbone conformation, residue interactions, residue contacts, and energy profile of the structure using different methods, including ERRAT [18], Verify 3D [19,20], and Molprobity [21]. The volume of the binding pocket was computed using the software of the CASTp server [22] with default settings.

3.3 Molecular docking

AutoDock Vina was used to perform the molecular docking analyses [23]. The model structures of RaUrease and its known inhibitor AHA were determined first. The molecules were prepared as docking input files using the AutoDock Tools 1.5.6 [24]. The amino acids surrounding AHA in the binding site of RaUrease were predicted. After that, a grid box with the dimensions 126 Å × 126 Å × 126 Å was built to cover the entire model structure of RaUrease. The grid center was set at 128.255 Å × 127.264 Å × 91.45 Å to cover each active site, and a grid spacing of 0.375 Å (approximately one-fourth of the length of the carbon–carbon covalent bond) was used for the calculation of the

energy map.

3.4 Virtual screening

The 3D structure of the established model RaUrease was used as the target protein. The Natural Products Database of the ZINC database was used because of its open source and commercial availability [25]. All compounds (11247 compounds in total) were analyzed by removing all the inorganic counter ions, adding hydrogen atoms, deprotonating strong acids, protonating strong bases, and generating stereoisomers. All compounds were then subjected to the same docking procedure. Compounds with docking scores of <-9.8 kcal mol^{-1} were selected for further virtual screening.

4. Conclusions

In conclusion, a reliable and reasonable 3D structure of RaUrease was successfully built with homology modeling techniques and molecular dynamics simulation methods. Twelve commercially available compounds were identified as potential urease inhibitors, and the binding site of the most effective compound on RaUrease was characterized. The findings of the present study will be of value to facilitate further experimental verification and to reduce nitrogen pollution derived from ruminant production. In addition, the present findings should help the design of novel inhibitors for enteric bacterial urease.

Acknowledgments： This study was financially supported by the Agricultural Science and Technology Innovation Program (Project Number: ASTIP-ISA12).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kertz, A.F. Review: urea feeding to dairy cattle: a historical perspective and review. **2010**, *26*, 257-272.
2. Recktenwald, E.B.; Ross, D.A.; Fessenden, S.W.; Wall, C.J.; Van Amburgh, M.E. Urea-N recycling in lactating dairy cows fed diets with 2 different levels of dietary crude protein and starch with or without monensin. **2014**, *97*, 1611-1622, doi:10.3168/jds.2013-7162.
3. Khattab, I.M.; Salem, A.Z.M.; Abdel-Wahed, A.M.; Kewan, K.Z. Effects of urea supplementation on nutrient digestibility, nitrogen utilisation and rumen fermentation in sheep fed diets containing dates. **2013**, *155*, 223-229, doi:10.1016/j.livsci.2013.05.024.
4. Jones, G.A.; Milligan, J.D. Influence on some rumen and blood parameters of feeding acetohydroxamic acid in a urea-containing ration for lambs. **1975**, *55*, 39-47.
5. Chan, C.C.L.; Jones, G.A. Effect of acetohydroxamic acid on growth and volatile fatty acid production by rumen bacteria. **1973**, *19*, 27-33.
6. Ludden, P.A.; Harmon, D.L.; Larson, B.T.; Axe, D.E. Influence of the novel urease inhibitor N-(n-butyl) thiophosphoric triamide on ruminant nitrogen metabolism: I. In vitro urea kinetics and substrate digestion. **2000**, *78*, 181-187.
7. Whitelaw, F.G.; Milne, J.S.; Wright, S.A. Urease EC 3.5.1.5 inhibition in the sheep rumen and its effect on urea and nitrogen metabolism. **1991**, *66*, 209-226, doi:10.1079/bjn19910026.
8. Hristov, A.N.; Oh, J.; Giallongo, F.; Frederick, T.W.; Harper, M.T.; Weeks, H.L.; Branco, A.F.; Moate, P.J.; Deighton, M.H.; Williams, S.R., et al. An inhibitor persistently decreased enteric methane emission from dairy cows with no negative effect on milk production. **2015**, *112*, 10663-10668, doi:10.1073/pnas.1504124112.
9. Wojcikowski, M.; Zielenkiewicz, P.; Siedlecki, P. Open Drug Discovery Toolkit (ODDT): a new open-source player in the drug discovery field. *Journal of cheminformatics***2015**, *7*, 26, doi:10.1186/s13321-015-0078-2.
10. Wei, Y.; Li, J.; Chen, Z.; Wang, F.; Huang, W.; Hong, Z.; Lin, J. Multistage virtual screening and identification of novel HIV-1 protease inhibitors by integrating SVM, shape, pharmacophore and docking methods. *European journal of medicinal chemistry***2015**, *101*, 409-418, doi:10.1016/j.ejmech.2015.06.054.
11. Kim, J.N.; Henriksen, E.D.; Cann, I.K.; Mackie, R.I. Nitrogen utilization and metabolism in Ruminococcus albus 8. **2014**, *80*, 3095-3102, doi:10.1128/AEM.00029-14.
12. Saeed, A.; Imran, A.; Channar, P.A.; Shahid, M.; Mahmood, W.; Iqbal, J. 2-(Hetero(aryl)methylene)hydrazine-1-carbothioamides as Potent Urease Inhibitors. **2015**, *85*, 225-230, doi:10.1111/cbdd.12379.
13. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. **2014**, *47*, 5 6 1-

32, doi:10.1002/0471250953.bi0506s47.

14. Baker, D.; Sali, A. Protein structure prediction and structural genomics. **2001**, *294*, 93-96, doi:10.1126/science.1065659.
15. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R., et al. Clustal W and Clustal X version 2.0. **2007**, *23*, 2947-2948, doi:10.1093/bioinformatics/btm404.
16. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]***2014**, *47*, 5.6.1-32, doi:10.1002/0471250953.bi0506s47.
17. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.; Smith, J.C.; Kasson, P.M.; van der Spoel, D., et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. **2013**, *29*, 845-854, doi:10.1093/bioinformatics/btt055.
18. Colovos, C.; Yeates, T.O. Verification of protein structures: Patterns of nonbonded atomic interactions. **1993**, *2*, 1511-1519, doi:10.1002/pro.5560020916.
19. Luthy, R.; Bowie, J.U.; Eisenberg, D. Assessment of protein models with three-dimensional profiles. **1992**, *356*, 83-85.
20. Bowie, J.; Luthy, R.; Eisenberg, D. A method to identify protein sequences that fold into a known three-dimensional structure. **1991**, *253*, 164-170, doi:10.1126/science.1853201.
21. Chen, V.B.; Arendall, W.B.; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: all-atom structure validation for macromolecular crystallography. **2010**, *66*, 12-21, doi:10.1107/s0907444909042073.
22. Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. **2006**, *34*, W116-W118, doi:10.1093/nar/gkl282.
23. Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **2010**, *31*, 455-461, doi:10.1002/jcc.21334.
24. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **2009**, *30*, 2785-2791, doi:10.1002/jcc.21256.
25. Irwin, J.J.; Sterling, T.; Mysinger, M.M.; Bolstad, E.S.; Coleman, R.G. ZINC: a free tool to discover chemistry for biology. **2012**, *52*, 1757-1768, doi:10.1021/ci3001277.

Table 1

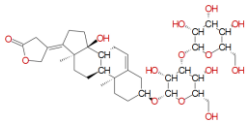
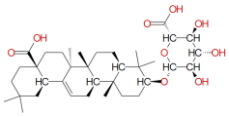
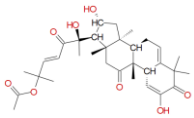
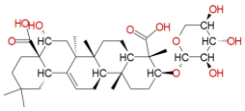
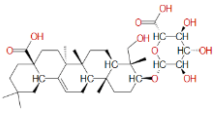
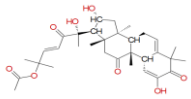
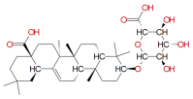
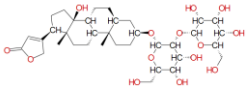
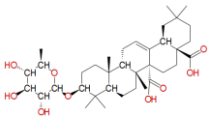
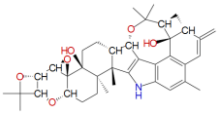
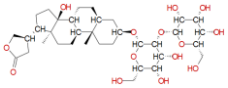
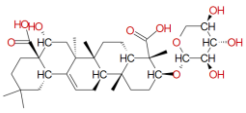
Validation of the three dimensional structures of model and template ureases

Validation	RaUrease (0 ns)	RaUrease (20 ns)	HpUrease
Favored regions	89.4% (508/568)	88.2% (500/567)	75.5% (426/564)
Allowed regions	97.5% (554/568)	98.6% (559/567)	92.7% (523/564)
Ramachandran outliers	2.46% (14/568)	1.41% (8/567)	7.27% (41/564)
Verify_3D	86.14%	89.28%	86.62%
ERRAT	82.740	90.556	85.536

The three dimensional structure of model urease from *Ruminococcus albus* 8 (RaUrease) was resolved (RaUrease, 0 ns) using the structure of urease from *Helicobacter pylori* (HpUrease) as template. The model structure was optimized using molecular dynamics simulation with a 20 ns procedure (RaUrease, 20 ns).

Table 2

Free binding energy among novel urease inhibitors and the *Ruminococcus albus* 8 urease

ZINC ID	Score (kcal mol^{-1})	Structure
ZINC67911797	-10.6	
ZINC67910936	-10.5	
ZINC59207168	-10.4	
ZINC67913308	-10.3	
ZINC67910929	-10.2	
ZINC59207170	-10.2	
ZINC67910942	-10.1	
ZINC67911804	-10.1	
ZINC67903153	-10.0	
ZINC67910930	-9.8	
ZINC67911581	-9.8	
ZINC69713310	-9.8	

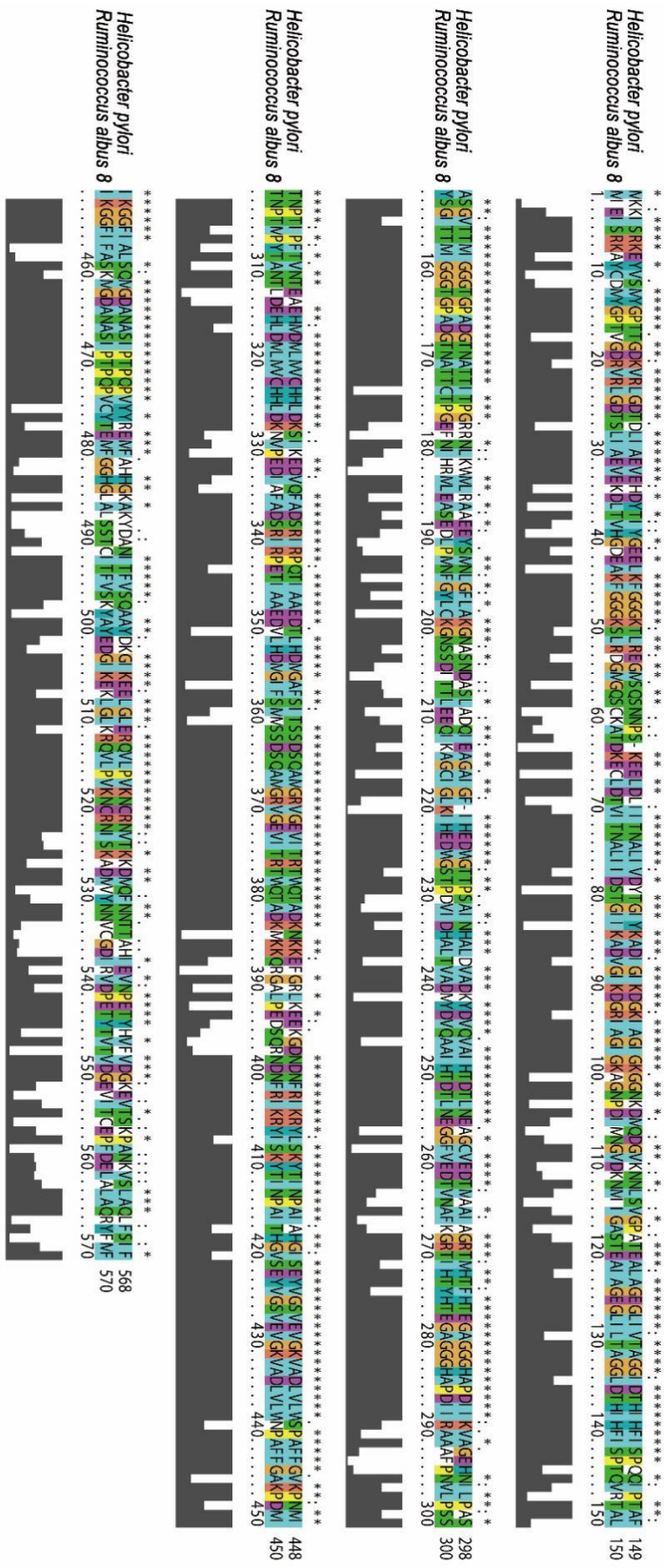


Figure. 1. Sequence alignment between ureases from *Ruminococcus albus* 8 and *Helicobacter pylori*. The sequence identity at amino acid level was 69%.

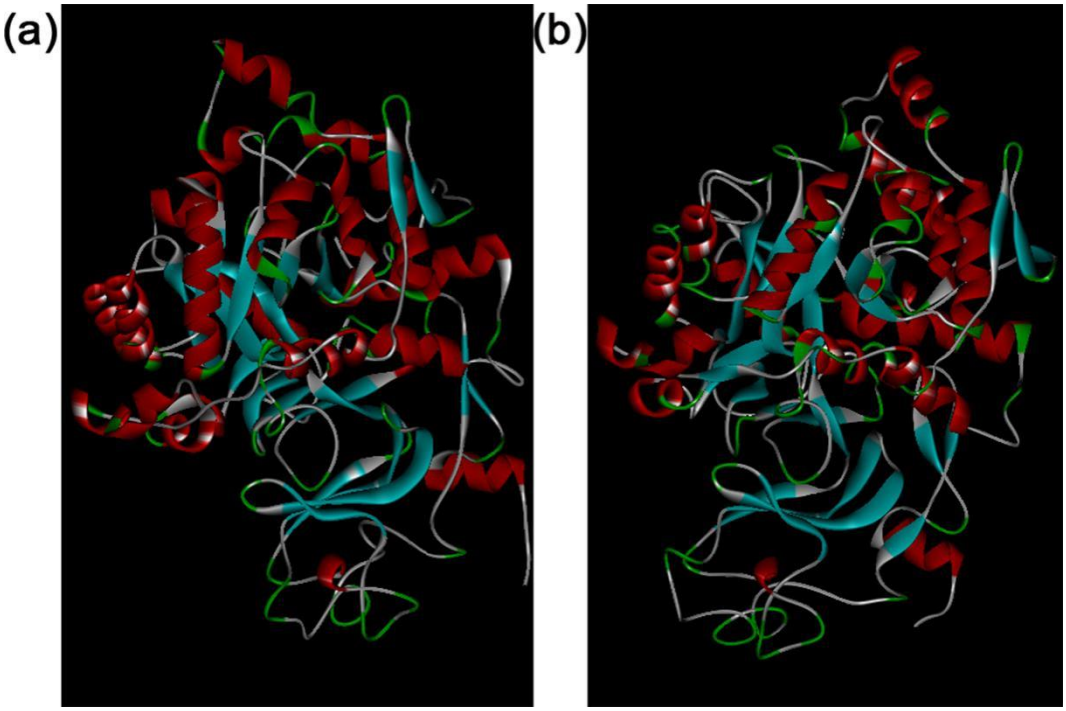


Figure 2. The predicted three dimensional (3D) model structure of *Ruminococcus albus* 8 urease was obtained using Modeller 9.16. The picture on the left (a) is the 3D structure established before molecular dynamics simulation, and the one on the right (b) is the structure after molecular dynamics simulation.

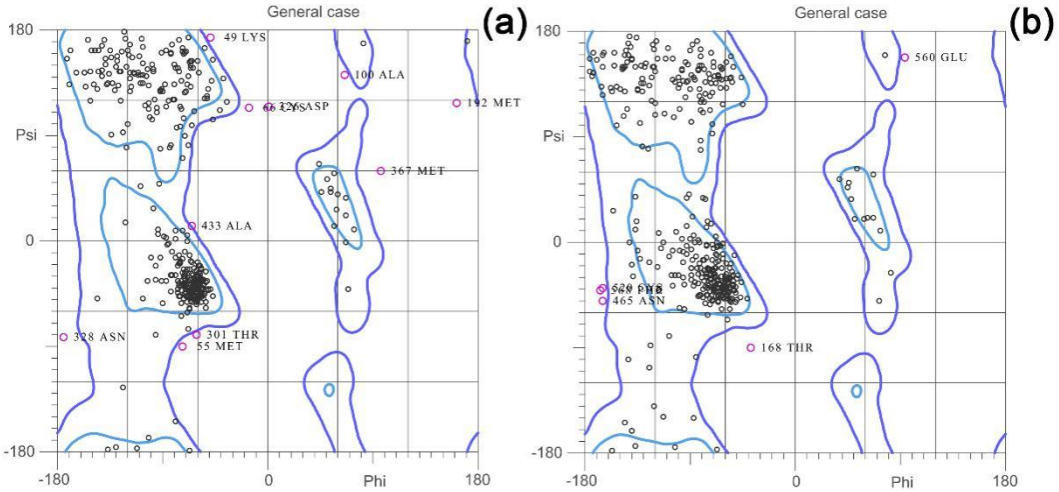


Figure 3. Ramachandran plots of *Ruminococcus albus* 8 urease initial (a) and final (b) three dimensional (3D) structures. There were 14 outliers (VAL17, LYS49, MET55, CYS66, ALA100, MET192, ILE287, THR301, ASN302, ASP326, ASN328, MET367, VaL427, ALA433) in the initial 3D structure of *R. albus* 8 urease, and eight outliers (ILE2, THR168, ASN302, PRO303, ASN465, CYS520, GLU560, PHE668) were still present after 20 ns of molecular dynamics simulation.

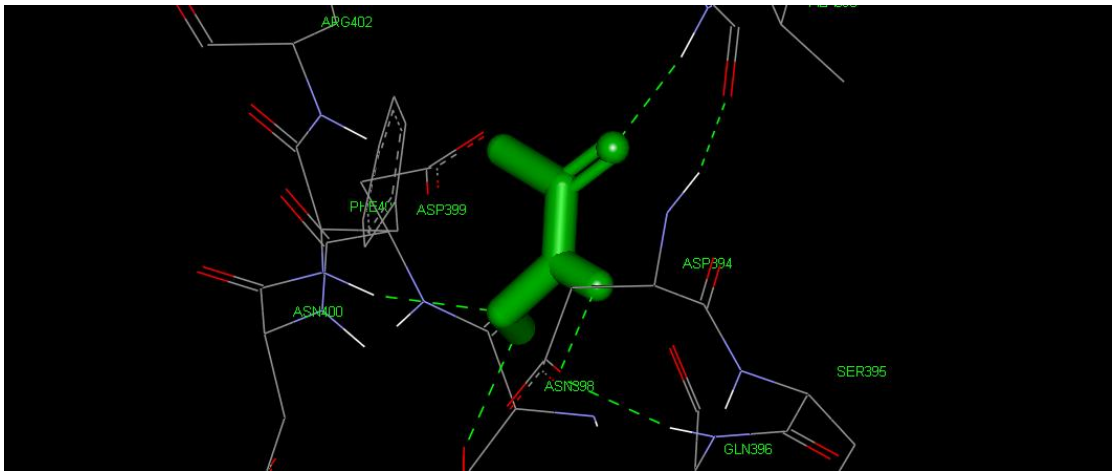


Figure 4. Predicted binding model of acetohydroxamic acid (AHA) to the model structure of *Ruminococcus albus* 8 urease. AHA was surrounded by ALA390, LEU391, ASP394, SER395, GLN396, ASN398, ASP399, ASN400, PHE401, and ARG402. There were five hydrogen bonds formed between AHA and the target protein.

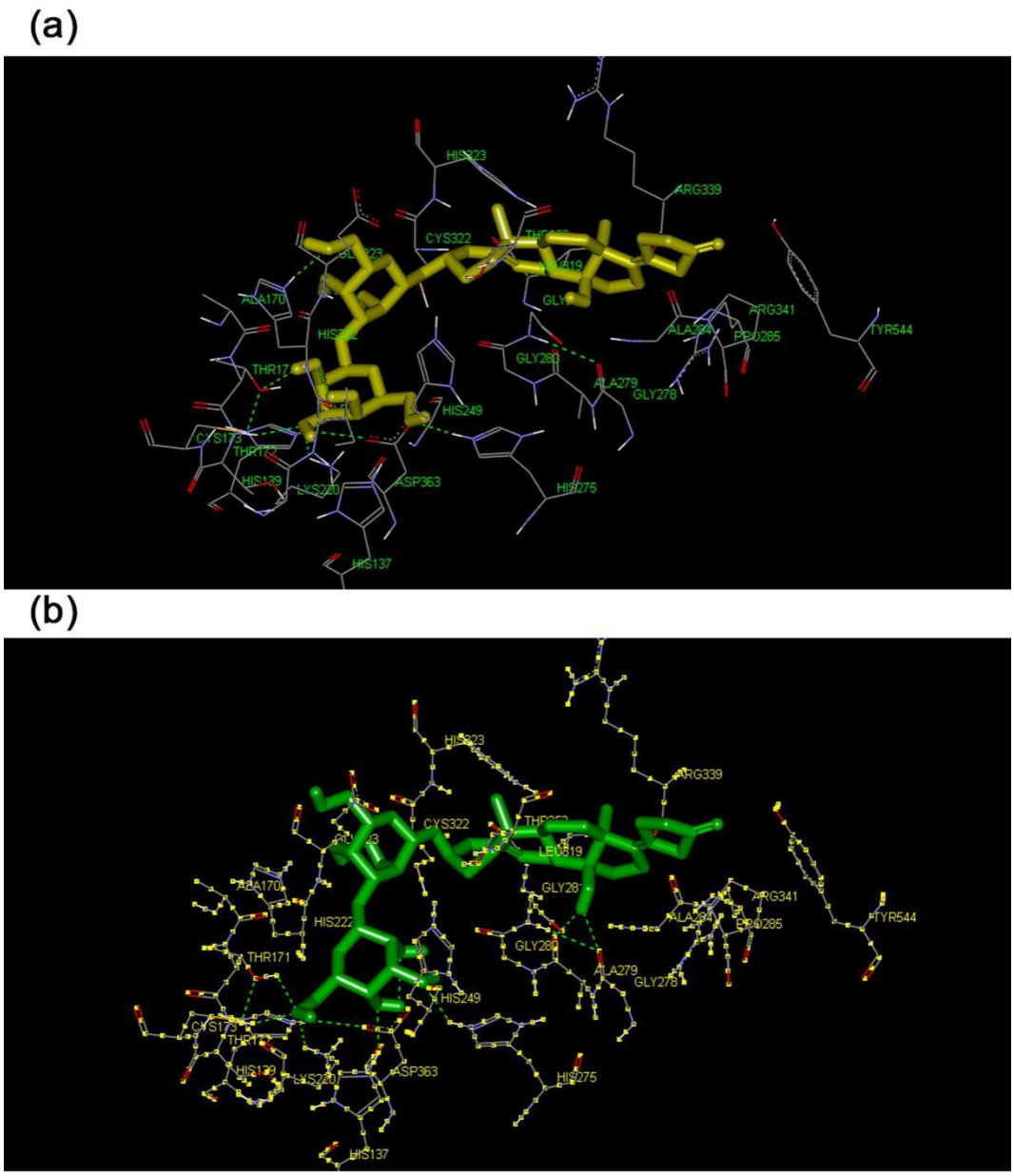


Figure 5. Predicted binding models of the best inhibitor ZINC67911797 (a) and its enantiomer ZINC67911804 (b) to the model structure of *Ruminococcus albus* 8 urease (RaUrease). The two inhibitors contained the same 26 amino acid residues (HIS137, HIS139, ALA170, THR171, THR172, CYS173, LYS220, HIS222, GLU223, HIS249, THR252, HIS275, GLY278, ALA284, PRO285, LEU319, CYS322, HIS323, ARG339, ARG341, ASP363, and TYR544), although ZINC67911797 had one more amino acid residue (ILE221) than ZINC67911804. There were five hydrogen bonds formed between ZINC67911797 and RaUrease, and seven hydrogen bonds formed between ZINC67911804 and RaUrease.