



Molecular modeling of *Helicobacter pylori* arginase and the inhibitor coordination interactions

Homa Azizian^{a,1}, Homayoon Bahrami^a, Parvin Pasalar^b, Massoud Amanlou^{a,*}

^a Department of Medicinal Chemistry, Faculty of Pharmacy, 16 Azar Ave., Tehran University of Medical Sciences, Tehran, Iran

^b Department of Biochemistry, Faculty of Medicine, Tehran University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Article history:

Received 22 September 2009

Received in revised form 1 December 2009

Accepted 21 December 2009

Available online 29 December 2009

Keywords:

Helicobacter pylori

Arginase

Docking

Molecular dynamics simulation

Homology modeling

ABSTRACT

Arginase of the *Helicobacter pylori* hydrolyzes L-arginine to L-ornithine and urea. *H. pylori* urease hydrolyzes urea to carbon dioxide and ammonium, which neutralizes acid. Both enzymes are involved in *H. pylori* nitrogen metabolism. The role of arginase in the physiology of *H. pylori* is metabolically upstream of urease which contributes in pathogenesis of this bacterium, so arginase could be potential drug target for *H. pylori* infection.

We performed homology modeling of *H. pylori* arginase using the crystal structure of *Bacillus caldovelox* arginase as a template, and then refined the model through molecular dynamics (MD) simulations.

Different criteria measured by PROCHECK, VERIFY-3D and PROSA were indicative of the proper fold for the predicted structural model of *H. pylori* arginase. Further evaluation on the model quality was performed by investigating the interaction of some arginase inhibitors with the modeled enzyme. Such interactions were determined employing Autodock 3.0.5 program. Our results are compatible with the published data on contribution of four aspartic acids: D116, D120, D234, D236 and three histidines: H91, H118, H133 for catalysis and stability of binuclear metal center of arginase that have important role in binding and catalytic activity in active site. In the absence of the experimental structure of *H. pylori* arginase we hope that our model will be useful to provide rational design of novel anti-*H. pylori* drugs.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Helicobacter pylori causes gastritis [1,2], and it is strongly associated with the development of peptic ulcers [3,4], and constitutes a risk factor for gastric adenocarcinoma [5]. In situ studies have shown that *H. pylori* synthesizes urea and ornithine from the catabolism of arginine by arginase [6], an enzyme of *H. pylori* urea cycle [7]. Then this urea is consumed by urease in order to catalyze it into carbon dioxide and ammonia. Consequently, the produced ammonia is essential for *H. pylori* to be protected from acid medium and to be colonized in the stomach, since *H. pylori* is an acid-sensitive organism [8,9]. Thus, in the metabolism of nitrogenous end products in *H. pylori*, arginase activity is upstream of urease [7], and as a result arginase could be a potential target to limit the source of urea and consequently ammonia, so inhibition

of the arginase has the ability to block the process of *H. pylori* colonization of stomach.

Arginase (L-arginine amidinohydrolase, E.C. 3.5.3.1) is a metalloenzyme and has two manganese ions that contribute in the hydrolysis of guanidine moiety of arginine. The proposed reaction mechanism involves a nucleophilic attack of a bridging OH⁻ ion on the guanidinium group of arginine, resulting in the formation of a tetrahedral intermediate [10]. Also arginase inhibits the production of nitric oxide, powerful oxidation species, via several potential mechanisms [11].

According to a comparison of *H. pylori* arginase amino acid sequence [12] with other sources of arginase sequence, the degree of identity and similarity between *H. pylori* arginase and those of other organisms was relatively low [13]. However it has been reported that there are seven important residues including; H91, H118, H133, D116, D120, D234 and D236, which are conserved and are shown to be important for catalysis and stability of binuclear metal center [14]. Additionally it has been revealed that the region 153–165 appeared as an insertion in the *H. pylori* sequence and was not found in the alignments with other arginase sequences. A possible role of this region is to interact with the membrane, thus resulting in the enzyme and cell envelope association [6]. Also it

* Corresponding author. Tel.: +98 21 66959067; fax: +98 21 66461178.

E-mail address: amanlou@tums.ac.ir (M. Amanlou).

¹ This research is a partial fulfillment of a PhD thesis in Medicinal Chemistry by Dr. Homa Azizian.

has been suggested that the N-terminal residues which do not include acidic residues, are probably involved in the mitochondrial localization of the enzyme [13].

In the absence of an experimentally determined crystal structure, homology modeling could provide a rational opportunity to obtain a reasonable 3D model. It is generally recognized that homology modeling of proteins is currently the most accurate method for 3D structure prediction, yielding models suitable for a wide spectrum of application, such as structure based molecular design and mechanism investigation [15]. This approach is able to provide a reasonable structure model with related template sharing more than 25% sequence identity. The present study attempts to propose the first model of *H. pylori* arginase. First, the 3D structure of the *Bacillus caldovelox* (PDB ID: 1CEV) [16] variant has been chosen as a template and then geometry of *H. pylori* arginase has been predicted using a homology modeling method. Also, the predicted structure has been refined by taking advantage of molecular dynamics (MD) simulation. Ultimately, in order to study the binding site, final geometry of *H. pylori* arginase was used in docking processes with the well-known ligands including arginine as a substrate and some inhibitors bond to the active site of arginase. NOHA [L-2-Amino-(5-(2'-hydroxyguanidino) penta-noic acid] is a competitive inhibitor of arginase with $K_i = 10\text{--}42\text{ }\mu\text{M}$ [17]. nor-NOHA [L-2-Amino-(4-(2'-hydroxyguanidino)butyric acid] which is the most potent analogue with $\text{IC}_{50} = 0.8\text{ }\mu\text{M}$, dinor-NOHA [L-2-Amino-(3-(2'-hydroxyguanidino) propanoic acid] and descarboxy-nor-NOHA are modified analogue of NOHA [18]. Also (s)-2-amino-7,7-dihydroxyheptanoic acid, (s)-2-(sulfonamidoethyl)-D-cysteine a sulfonamide transition state inhibitor [19] and a branched hydrophobic amino acids L-valine, $K_i = 3.6\text{ mM}$ [20], were used in order to investigate the validity of homology modeled protein.

2. Methodology

2.1. Amino acid sequence alignment

In homology modeling phase, we would like to look for an experimentally determined structure of high sequence identity with the *H. pylori* arginase. The amino acid sequence of the *H. pylori* arginase was retrieved from Swiss-Prot database (Primary accession number A3R4E2, wild type *H. pylori* strain 43504, Gene name RocF). The BLAST (version 2.2.18) search engine, which is publicly available at National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST/>) was used to find the homologous proteins with known structures to be used as the template in comparative modeling of arginase [21]. For the alignment between sequences of arginase and few homologous proteins, CLUSTALW was used from its web site at <http://www2.ebi.ac.uk/CLUSTALW> [22]. Among them *Bacillus caldovelox* arginase (PDB ID: 1CEV) [16] with known structure was used as the template for model building. The bacterial arginase is a homo-hexamer, so we extracted the monomer structure and used it as our structure template. There were 27.8% identity and 48% similarity between sequences of *H. pylori* and *B. caldovelox* arginase. Therefore, we used 1CEV as our template in the following homology modeling.

2.2. Construction of *H. pylori* arginase by homology modeling

The 3D-model of the *H. pylori* arginase apo-monomer was built based on template using MODELLER 9v4 [23]. The heteroatom module in the MODELLER 9v4 was used to model the Mn ions. Utilizing this program, fifty satisfactory models were generated for the *H. pylori* arginase. The best one according to the lowest MODELLER objective function and PROCHECK [24] Ramachandran

plot statistics was further refined by the energy minimization and MD simulation.

2.3. Refinement of homology model

The initial model was refined with MD simulation which was carried out with the parallel version of GROMACS 4.0.3 [25] in an all-atom force field (OPLS-AA). Before starting MD simulations, residues pK_a were calculated to determine if any of the residues in the enzyme were likely to adopt non-standard ionization states by PROPKA 2.0 [26]. Only His112 was completely protonated while Lys220 was deprotonated and other residues followed standard ionization state. Then the protein was put into a suitable sized simulation cubic box and solvated with 17,835 simple point charges (SPC) water model. Also 3 Na^+ counter ions were added to neutralize the negative charge. First, the water molecules and ions alone were subjected to energy minimization with the *H. pylori* arginase which was kept fixed in its initial configuration. The water molecules and ions were then allowed to evolve using MD simulation for 20 ps with a step time of 1 fs, again keeping the structure of the *H. pylori* arginase fixed. Next, the entire system was minimized using the steepest descent of 1000 steps followed by conjugate gradients of 9000 steps. In order to obtain equilibrium geometry at 300 K and 1 atm, the system was heated at a weak temperature ($\tau = 0.1\text{ ps}$) and pressure ($\tau = 0.5\text{ ps}$) coupling by taking advantage of the Berendsen algorithms [27]. Heating time for MD simulations at 100 K, 200 K was 100 ps. All above simulations were performed at constant temperature and pressure with a non-bonded cut-off of 14 Å. The MD simulation was further carried out for 30 ns at 300 K, followed by structural minimization calculation. The latter minimization was performed at the steepest descent of 1000 steps followed by conjugate gradients of 9000 steps. LINCS was used to constrain bond length. The MD simulations were carried out employing particle mesh Ewald method [28] for the electrostatic interactions. The integration time step was 1 fs, with the neighbor list being updated every fifth step by using the grid option and a cut-off distance of 12 Å. Periodic boundary condition has been used with constant number of particles in the systems, pressure, and temperature simulation criteria (NPT). During the simulation every 1.000 ps of the actual frame was stored. The energy minimized average structure calculated from the equilibrated trajectory system was evaluated for quality determination of protein geometry and structure folding reliability. Subsequently, the dynamic behavior and structural changes of the protein was analyzed by the calculation of energy and the root mean square deviation (RMSD).

2.4. Assessment of the homology model

The quality of protein geometry was checked employing PROCHECK (version 3.5.4) [24], Verify-3D used from its web server (<http://www.doe-mbi.ucla.edu/Verify3d.html>) [29] and ProSA [30]. The PROCHECK is applied to qualify the combinations of ϕ and ψ residues angles in available zones of Ramachandran plot while Verify-3D uses a score function to assess the quality of the model and ProSA program evaluates the energy of the structure using a distance based pair potential. Also RMSD values between superimposed model and crystallographic structure were calculated by using SuperPose program [31].

2.5. Validation of the model by docking analysis

Autodock 3.0.5 [32], which uses a stochastic-search based algorithm, was employed for the generation of ligand–protein complexes. AutoDockTools (ADT) was used to add hydrogen, charges and active torsions to the ligands and also it was used to



Fig. 1. Sequence alignment of *H. pylori* arginase with *B. caldovelox* (PDB code: 1CEV) arginase generated by CLUSTALW algorithm. An asterisk (*) indicates an identical or conserved residue, a colon (:) indicates a conserved substitution, a stop (.) indicates a semi-conserved substitution. Secondary element of modeled structure using DSSP is also shown. Alpha helices, beta sheets are shown in gray, and yellow, respectively. The active site amino acid residues corresponding to the active site of 1CEV derived from crystallographic data are shown in bold red color.

assign Kollman charges to the enzyme. As our purpose was to assess the binding site of arginine we made the search space large enough that included the whole enzyme and increased exhaustiveness with the affinity grids of $126 \times 126 \times 126$ points 0.375 \AA° around the protein using AutoGrid. The Lamarckian Genetic Algorithm (LGA) was used for the conformational search. Each Lamarckian job consisted of 250 run. The initial population was 150 structures, and the maximum number of energy evaluations and generations was 10,000,000. The rest of the parameters were set to default values. These values were chosen based on the systematic study on the influence of the different parameters in dockings without prior knowledge of the binding site [33]. The final structures were clustered and ranked according to the native Autodock scoring function. The top-ranked conformation of each ligand was selected.

In the current study the binding modes of arginine as a substrate and NOHA, nor-NOHA, dinor-NOHA, descarboxy-nor-NOHA, (s)-2-amino-7,7-dihydroxyheptanoic acid, (s)-2-(sulfonaminoethyl)-D-cysteine and L-valine as selective competitive inhibitors were the case in to the whole protein in order to investigate the validity of modeled enzyme.

2.6. Computational resources

The computational studies were carried out on computer cluster of 4 HP Proliant ML370-G5 Tower servers equipped with 2 Quad Core Intel Xeon processors E5355, 2.66 GHz and 4 GB of RAM on a Linux platform.

3. Result and discussion

3.1. Sequence alignment and model building

Fig. 1 shows the sequence alignment between 1CEV and *H. pylori* arginase that is needed to perform homology modeling. MODELLER 9v4 was used to generate model structure based on *B. caldovelox* arginase (PDB; 1CEV) as template. Since it has been

suggested that the N-terminal residues which involved in the mitochondrial localization of the enzyme [13], we built one monomer of the arginase from residue 60–284 in the α/β domain.

The monomer of the modeled target is folded into α/β domain and is comprised of five-stranded parallel β -sheets with helices packed on either side. The strands of the central β -sheet are arranged in the order β_1 – β_5 – β_4 – β_2 – β_3 . One or two α -helices connect one strand to the next in such a way that a set of four helices (α_1 – α_2 – α_3 – α_4) covers one face of the β -sheet and the other set of helices (α_5 and α_6) covers the other face (Fig. 2a). The same α/β fold is shared by other arginase family enzymes include *B. caldovelox* arginase [16].

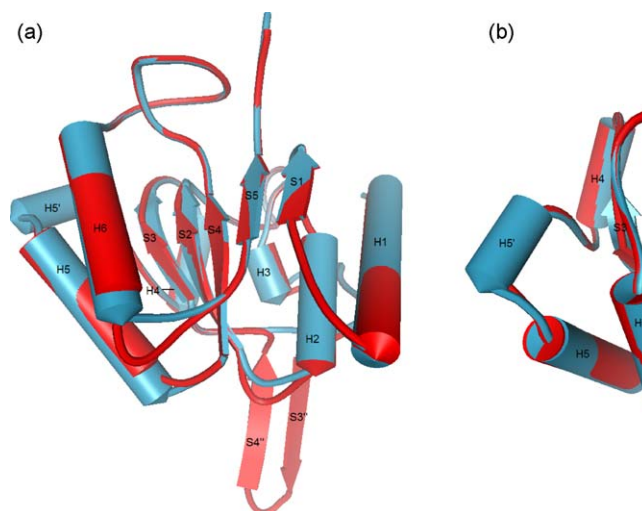


Fig. 2. Comparison of the *H. pylori* homology model and X-ray structure. (a) Superposition of the model and the template structures after free MD simulation. Helices are depicted by cylinders, β -sheets by arrows, random coil by tube. *H. pylori* and 1CEV are colored red and blue, respectively. (b) Detailed view of the region where model and template show deviation.

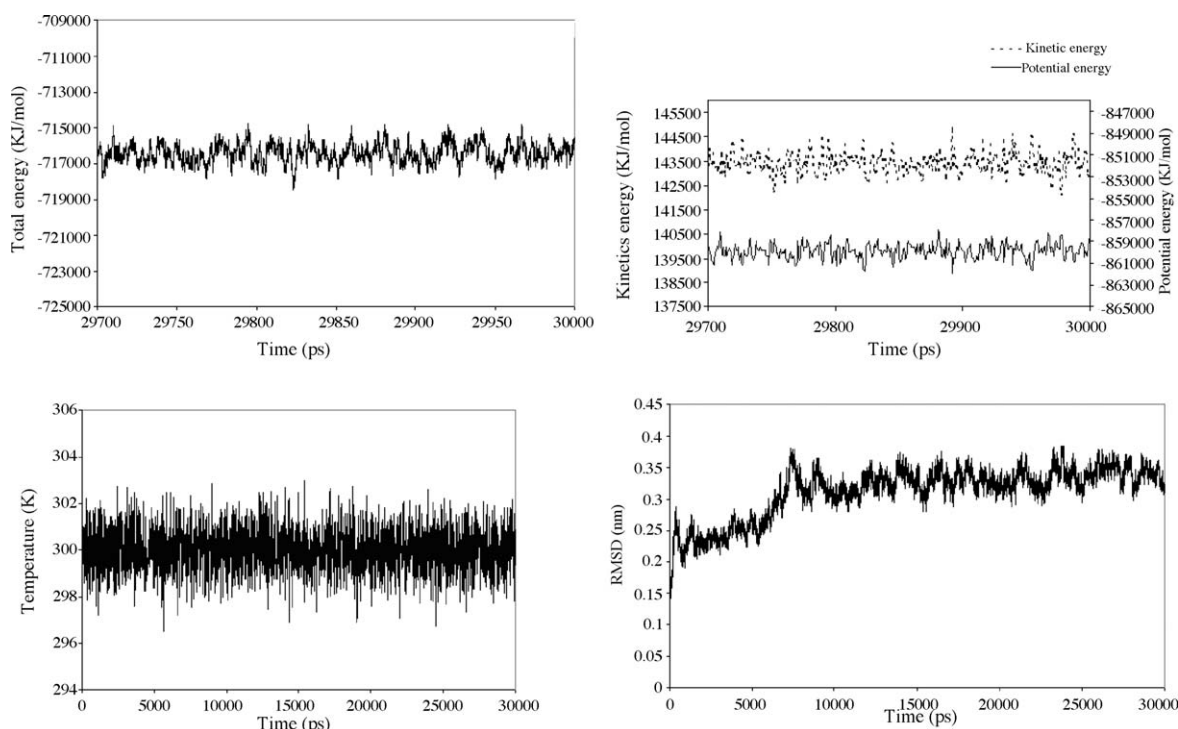


Fig. 3. The variation of the total energy during the last 300 ps of molecular simulation (a). The kinetics and potential energy components fluctuation in equal and opposite direction during the last 300 ps of simulation (b). Backbone root mean square deviation (RMSD) of the model during the simulation (c). The variation of the temperature during molecular simulation (d).

To improve and verify the stability of initial MODELLER structure, MD calculation was carried for 30 ns. Variation in total energy versus time in last 300 ps of MD simulation is shown in Fig. 3a. It shows small fluctuation through this time. Also, by taking a look at Fig. 3b, it reveals that the kinetics and potential energy components in last 300 ps of simulation would be expected to fluctuate in equal and opposite direction. These facts show that energy conservation satisfies in MD simulation performed [34]. It is noteworthy that the RMSD of structures relative to the starting structures were $\sim 4 \text{ \AA}$ after 15 ns of simulation, and the RMSD values did not changed significantly after 15 ns of simulations as it is illustrated in Fig. 3c. These RMSD values indicate that the employed simulation time has been enough to obtain an equilibrium structure of the *H. pylori* arginase and basic variation occurred in latter geometry during 15 ns. Thus, the applied MD simulation was necessary to specify geometry of *H. pylori* arginase. In addition, average temperature of 30 ns of simulation at 300 K for the studied system was equal to $300 \pm 0.9 \text{ K}$ (Fig. 3d). Therefore, the extracted equilibrium structure at 300 K belonging to the *H. pylori* arginase was obtained under temperature stable conditions.

3.2. Quality assessment of the modeled arginase

The quality of the homology model was evaluated by PROCHECK, the detailed residue-by-residue stereochemical quali-

ty of the selected model, which was found to be good (83.1% in favored core regions, 15.7% in allowed regions comparing with the template, also with relatively low percentage of residue 0.6% in generously allowed regions and 0.6% in disallowed regions) and the total quality of G-factors was obtained to be 0.14 which indicates the designed models good quality (Table 1) (acceptable values of the G-factor in PROCHECK are between 0 and -0.5 , with the best model displaying values close to zero). Residues N211 and K109 are in disallowed and generally allowed region, respectively (Fig. 4). These positions do not interfere with catalytic site of arginase, since their locations (blue colored in Fig. 5) are far from the active site (magenta colored).

Verify-3D uses a score function to analyze the compatibility of the residues with their environment in MODELLER result and the final MD model. Residues with a score over 0.2 should be considered reliable. It can be seen from the Fig. 6 that the score of refined structure after MD simulation (blue line) is clearly better than the one from MODELLER (magenta line), and also there is no residue with negative compatibility score after MD simulation which indicates the proper fold of model. 88.8% of the residues have a score over 0.2. Fig. 6 also shows that the sequence between residues 152 and 166 exhibits lower scores. This fragment corresponds to a sequence of the *H. pylori* arginase that includes one gap in the alignment that is correspondent to the inserted fragment (Fig. 1) [11]. However, this fragment is solvent exposed,

Table 1
Quality of structures checked by PROCHECK for target and template.

PROCHECK ^a	Ramachandran plot quality (%)				Goodness factor		
	Core	Allowed	General	Disallowed	Dihedral	Covalent	Overall
Model	83.1	15.5	0.6	0.6	-0.19	-0.1	-0.14
Template	91.2	8.8	0	0	0.19	0.44	0.29

^a Ramachandran plot qualities show the percentage (%) of residues belonging to the core, allowed, generally allowed and disallowed region of the plot; goodness factors shows the quality of covalent and overall bond/angle distance; these scores should be above -0.5 for a reliable model.

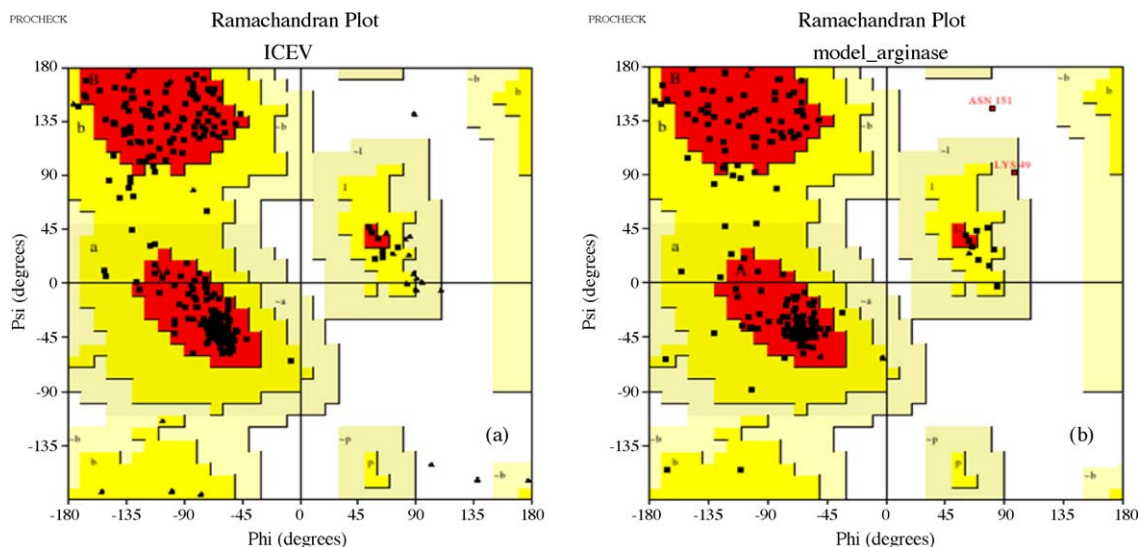


Fig. 4. Ramachandran plots of modeled arginase (a), and the *B. caldovelox* crystal structure (b). The favored and most favored region is yellow and red respectively, pale yellow is the generally allowed and disallowed regions is white.

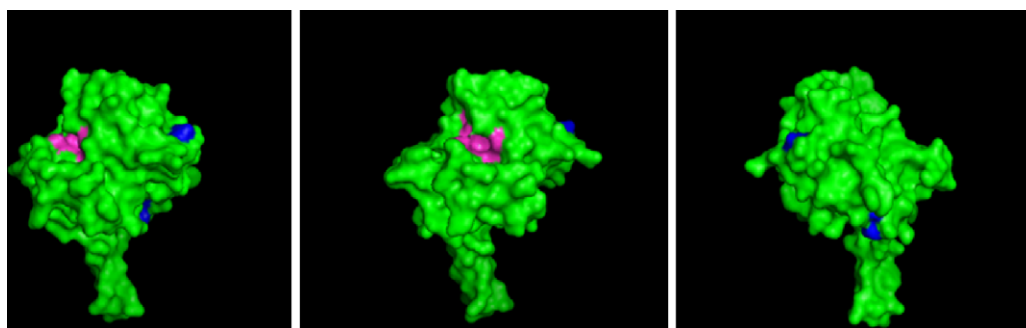


Fig. 5. Surface representation of modeled *H. pylori* arginase. Magenta region is related to active site residues. Residues in disallowed region and general region of the Ramachandran plot are shown in blue. The model is twisted 90 and 180 degree in middle and right picture.

with no role in stability. Therefore in the rest of study we do not intend to focus on this modeled fragment.

Finally, Fig. 7 shows the energy profiles for MODELLER model and for the final MD structure using the ProSA score. This program evaluates the interaction energy per residue of the structure using a distance based pair potential. Residues with negative ProSA energies confirm the reliability of the model. In our model the positive score residues belong to the inserted fragment in the *H. pylori* arginase which has no template for modeling. The ProSA result obtained from MD simulation has a better score than

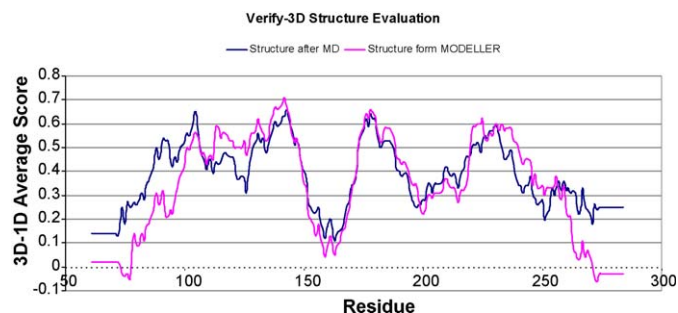


Fig. 6. Verified-3D analysis is shown for residue between 60 and 289; the positive scores suggest that the residues are compatible with their environments in the model build for *H. pylori* arginase.

MODELLER by decreasing the number of residues as well as values with positive peaks (Fig. 7). To sum up, the geometric quality of the backbone conformation, the residue interactions and the energy profile of the structure are all well within the limits established for reliable structures.

To investigate how well the modeled structure matches the X-ray data of template, the *H. pylori* arginase homology model and crystal structure of 1CEV were superimposed on their backbone atoms (Fig. 2a). RMSD values for the backbone atoms were 3.63. However, from a visual inspection a good overall agreement of secondary structural elements of the homology model and the X-ray structures is observed (Figs. 1 and 2a). In fact, high RMSD values originate mainly from large deviations in the H4–S3 region. Here the model possesses 2 extra β -sheets (S3'' and S4'') which are related to the special inserted sequence in *H. pylori* arginase. Excluding this segment (amino acids 157–172) from the measurement, the RMSD dropped significantly to value 0.74 for 204 carbon- α atoms and 0.78 for 816 backbone atoms. According to the reports these residues do not influence arginase catalytic activity and they may involve in the association of enzyme with cell envelop [6]. Additionally, 1CEV has an additional helix (H5') which connects H4 and H5, whereas the homology model contains a flexible loop with the same orientation (Fig. 2b). This difference is because of gap in this region in model sequence.

To further evaluate the accuracy of the active site model we compared the side chain orientations of the amino acids distances

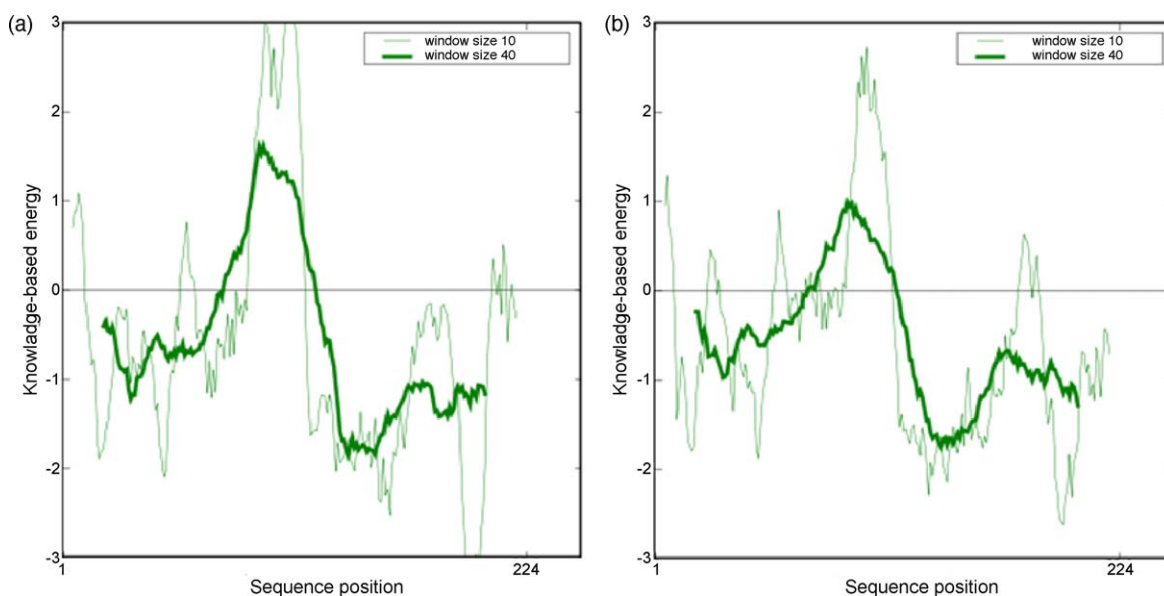


Fig. 7. ProSA energy profiles calculated for the MODELLER model (a) and the final molecular dynamic structure (b).

8 Å from Mn^{2+} ions in the model and the X-ray structure. The active site of model and X-ray structure were superimposed and the RMSD of all atoms was 0.131. Fig. 8 shows the superimposed binding pockets of the homology model and the X-ray structure 1CEV illustrating the high accuracy of the *H. pylori* active site model.

All evaluations suggest that a reasonable homology model for the *H. pylori* arginase monomer has been obtained in order to examine the protein–substrate interactions. Ideally, it is better to run the hexamer (tri-dimers) arginase simulation to well understand the inter monomer dynamics. However it would be expensive in terms of computing cost. Our major goal is to characterize the residue involving in active site and substrate binding mode also we know that the whole active site is located in each monomer and each of them has separate catalytical activity. Therefore, modeling of the monomer rather than oligomer will not change our major conclusion of this study.

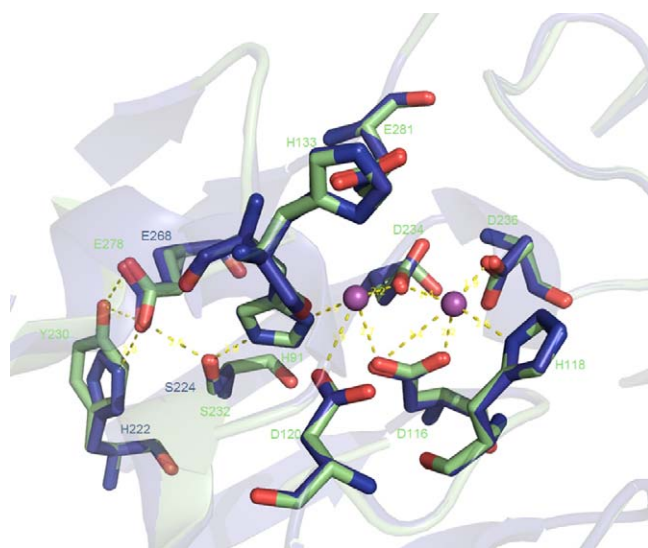


Fig. 8. Structural alignment of target and template active site. Residues of 1CEV are colored in blue, whereas amino acids of the homology model are shown in green.

3.3. Study of conserve residues

We used our homology modeled structure to search for conserved residues in order to obtain the corresponding findings for *H. pylori* arginase. Analysis of the sequence conservation between *H. pylori* arginase and 1CEV identified by alignment of sequence data reveals striking conservation involving the six metal-binding ligands; H99, D122, H124, D126, D226 and D228 (using the *B. caldovelox* numbering). For the *H. pylori* arginase the six metal-binding residues are H91, D116, H118, D120, D234 and D236.

The active site of *H. pylori* arginase is shown in Fig. 10. The Mn^{2+} – Mn^{2+} separation is 3.4 Å, identical to that in 1CEV as template [16]. One of the Mn^{2+} ions is coordinated by D120 (OD2), D116 (OD2), D116 (OD1), H91 (ND1), and D234 (OD2) and the other Mn^{2+} ion is coordinated by D234 (OD1), D236 (OD2), H118 (ND1) and D116 (OD1). Comparison of the X-ray crystallographic structures of 31 arginase family enzymes [35] led to the conclusion that Mn^{2+} -binding sites in enzymes favor of highly conserve histidine and aspartic acid as the principal ligands involved in binding and catalysis. Based on the experiment and our theoretical predicted results, in the present study this site is chosen as the more favorable binding site to dock the ligands.

Also seen was the significant conservation of the S232 residue (S224 in the *B. caldovelox*). Its hydroxyl group lies 2.8 Å from NE2 atom of H91, whose ND1 atom coordinates one of the two Mn^{2+} ions. It was reported that arginase activity in extracts with the S232I mutation was abolished [36]. Additionally two side chain oxygens of highly conserved D278 make hydrogen bond with hydroxyl group of S232 (2.6 Å) and Y230 (2.9 Å and 3.0 Å) (Fig. 8). S232 and D278 are indirect metal-binding residues, which through hydrogen binding to direct metal ligands, affects the overall electrostatic potential of the metal-binding site [37]. Also Y230 is occupied by H222 in *B. caldovelox* (Fig. 1). H222 is conserved in most arginases and it is replaced by aromatic residue in arginase related enzymes like agmatinases and formiminoglutamase [35].

E281 is also another conserved residue which equivalent with E271 in *B. caldovelox*. It may involve in forming salt bridge with the charged nitrogen in the substrate. Many available arginase family enzyme sequences reveal this residue could help position of guanido group of substrate for attack by nucleophile [38].

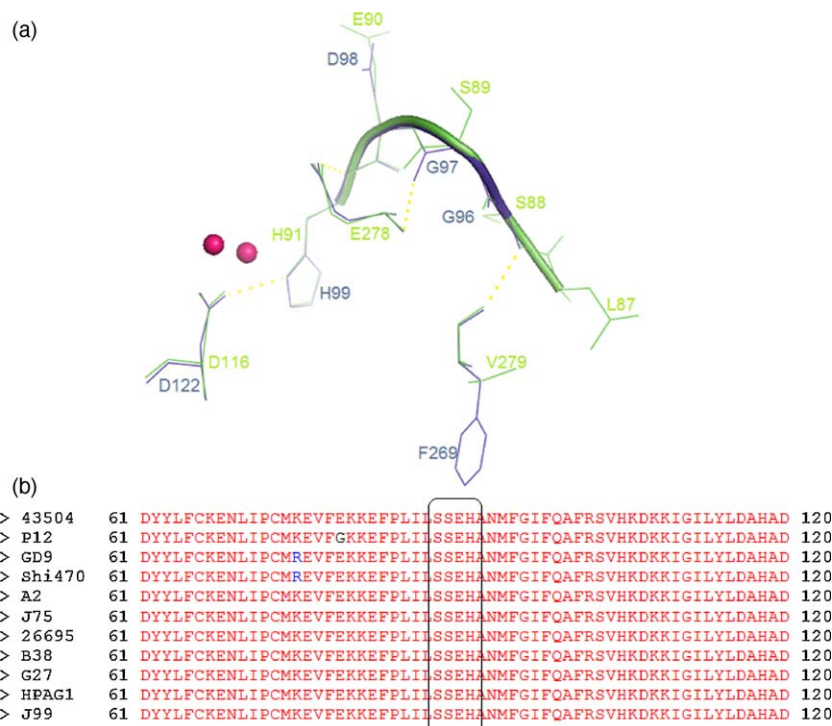


Fig. 9. 3D representation of GGDH and SSDH motif in 1CEV (blue) and *H. pylori* arginase (green) (a). Sequence alignment of ten *H. pylori* strains which obtained from Swiss-Prot database, 43504 (SWISS-PROT accession number A3R4E2); P12 (SWISS-PROT accession number B6JP31); GD9 (SWISS-PROT accession number A3R4G1); shi470 (SWISS-PROT accession number B2UVT6); A2 (SWISS-PROT accession number A3R4F6); J75 (SWISS-PROT accession number A3R4E5); 26695 (SWISS-PROT accession number O25949); B38 (SWISS-PROT accession number C7C183); G27 (SWISS-PROT accession number A3R4G3); HPAG1 (SWISS-PROT accession number Q1CR85); J99 (SWISS-PROT accession number Q9ZD87) (b).

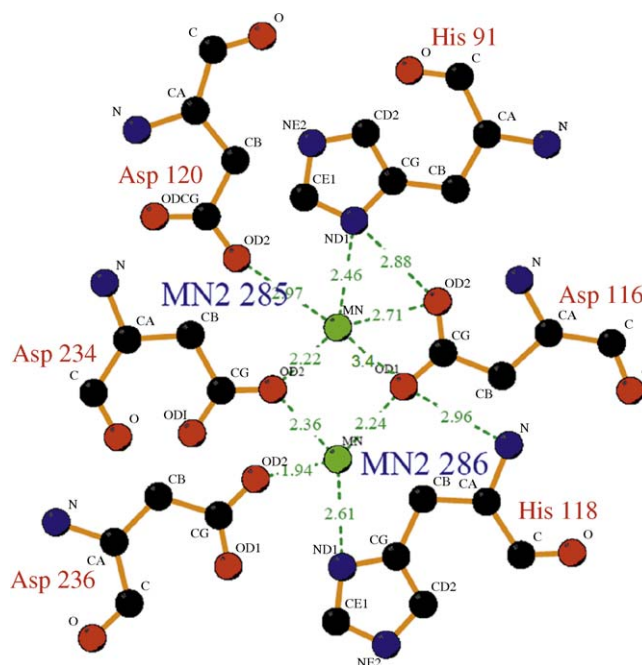
B. caldovelox arginase has a *cis*-peptide bond between G96 and G97 which is part of the conserved GGDH motif in the arginase superfamily [39]. The backbone nitrogen atom of G96 is hydrogen-bonded to the carbonyl oxygen of F269, and the carbonyl oxygen of G97 is hydrogen-bonded to the backbone nitrogen of E271 (Fig. 9a). This network of hydrogen bonds is important in positioning the key active site residues H99 and E271. Although this motif was changed to SSDH in *H. pylori* arginase, the *cis*-peptide was configured because of the hydrogen-bonded of backbone nitrogen of S88 to the carbonyl oxygen of V279, and the carbonyl oxygen of S89 to the backbone nitrogen of E278 (Fig. 9a). This event is supposed to the small character of serine that could be substitute of glycine in this motif. In order to examine this conserve motif in other *H. pylori* arginase strains, we have done sequence alignment with 10 different strains of *H. pylori* arginases. So we did sequence alignment with ten different strains of *H. pylori* arginases. It is obvious from Fig. 9b that the SSDH motif can be find in other strains of *H. pylori* arginases, while this motif is GGDH in other arginases family.

Finally there are four hydrophobic similarities between target and template, respectively (I112/V118, L113/I119, L235/L227, and L266/L258), which appear necessary for packing within the hydrophobic protein interior [35].

3.4. Molecular docking

The ability to predict the correct ligand binding pose is important for a successful homology modeling and the main validation is to check whether it is in consistent with experimentally or published data. To further explore the common characteristic of the *H. pylori* arginase binding site the arginine and some inhibitors were docked in to the whole enzyme with the active conformation after MD simulation. The reliability of the applied docking protocol was assessed by docking arginine into the active

site of the modeled *H. pylori* arginase (Fig. 10). The key characteristic of a good docking program is its ability to reproduce the experimental binding modes of ligands. To test this, a ligand is taken out of the X-ray structure of its protein–ligand complex and docked back into its binding site. The docked binding mode is then compared with the experimental binding mode, and RMSD is calculated; a prediction of a binding mode is considered successful



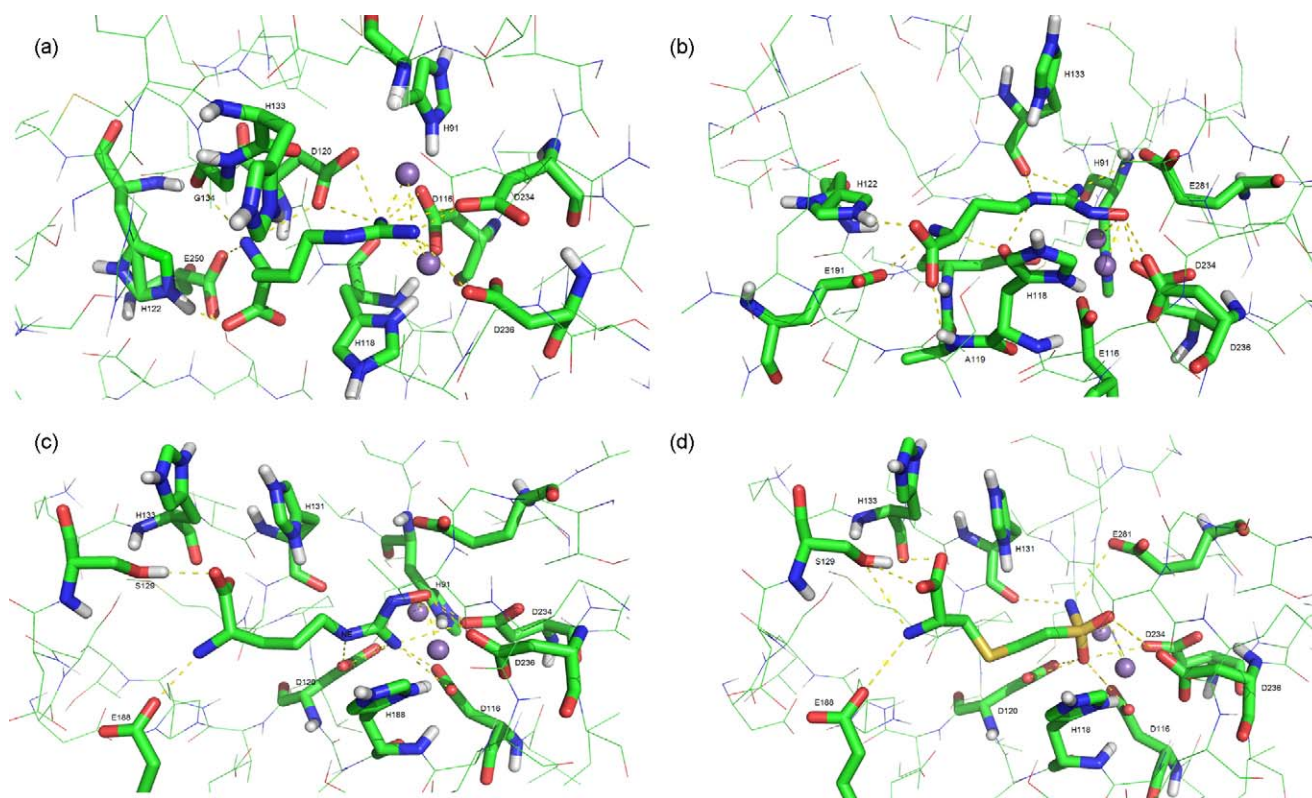


Fig. 11. 3D representation of arginine (a), NOHA (b), nor-NOHA (c) and (s)-2-(aminosulfonyl)-D-cysteine (d) in the active site after docking.

if the RMSD is below a certain value (usually $< 2.0 \text{ \AA}$). The RMSD between the docked binding mode and the experimental binding mode for arginine when docked into the modeled structure was within this cut-off limit (1.42 \AA). This protocol was then similarly applied to all ligands.

Fig. 11 represents orientation and conformation of arginine, NOHA, nor-NOHA, and (s)-2-(sulfonamidoethyl)-D-cysteine in the binding site. Also Table 2 shows the structures and related docking energies. From Fig. 11a it is obvious that the guanidine moiety has an important role for the orientation of arginine in its active site which coordinates with D236 (OD2), D234 (OD2), D120 (OD1, OD2) and D116 (OD2) and imposed ionic interaction with 2 Mn^{2+} ions. About 81% of all runs construct this conformation. Additionally, α -amino moiety of arginine has polar contact with D120 (OD1), G134 (O), E250 (OE2), while its α -carboxylate participates hydrogen bond with NE2 atom H122 (1.8 \AA).

In the case of NOHA, about 70% of all runs orient in the manner that the hydroxyl group bridge the binuclear manganese cluster, occupying the site normally populated by the bridging hydroxyl group in the free enzyme, and also it coordinates by D234 (OD2), D236 (OD2) with distance about 2.5 \AA and 2.9 \AA respectively. The NH_2 group of hydroxyl-guanidine moiety coordinates by Mn^{2+} and donate hydrogen bond with D234 (OD2), D120 (OD1, OD2), D116 (OD1, OD2), which stabilized the tetrahedral transition state. Similarly the NE atom contributes in hydrogen bond with the backbone carbonyl group of D120. Also the α -carboxyl group and α -amino moiety of NOHA participates in hydrogen bonding with the hydroxyl group of S129 and OD1 atom E188, respectively (Fig. 11b).

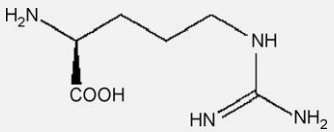
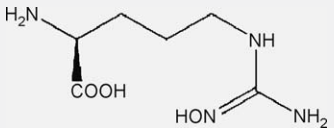
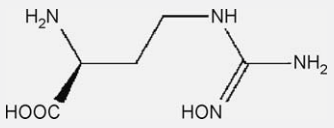
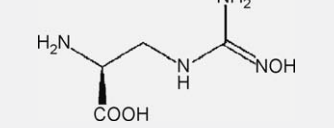
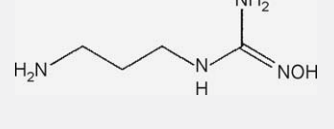
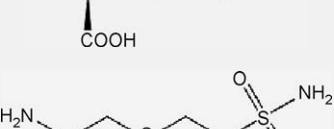
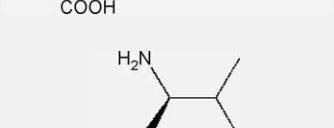
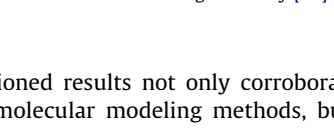
The nor-NOHA hydroxyl-guanidino group is bound to the Mn^{2+} ions. The interaction between polar groups of nor-NOHA and arginase is shown in Fig. 11c. Structure activity relationships for arginase analogues indicate that electrostatic interactions with α -substituent of the substrate are critical for catalysis [40]. The shorter length of nor-NOHA rather than NOHA could help it to

direct hydrogen bonds of the α -amino group with D120 and E191 and α -carboxyl group interact with NE2 atom H122 and backbone amid of A119. This finding was supported by lower docking energy of nor-NOHA rather than NOHA (-10.71 vs. -9.28) (Table 2).

Dinor-NOHA and descarboxy-nor-NOHA have the same conformation and orientation as NOHA in the active site with different docking energy (Table 2). This finding is compatible with the crystallographic data in the study of inhibitor coordination with recombinant rat liver arginase that hydroxyl-guanido group of inhibitors bridge the Mn^{2+} ions and α -amino and α -carboxyl group have important role in stabilizing of ligands in active site [41].

Docking of (s)-2-(sulfonamidoethyl)-D-cysteine to the modeled enzyme revealed that the NH_2 group of sulfonamide interacts with carbonyl group of H131 and OE1 atom of E281. Also the two oxygens bridge the binuclear manganese cluster with getting hydrogen bond with D236, D234 and D116. The α -amino group participates in hydrogen bond with E188 and S129, and the two oxygen atoms of α -carboxylate interact with S129. This result is similar to the rat liver arginase crystallography study by Christianson et al. that the α -amino group of (s)-2-(sulfonamidoethyl)-D-cysteine interacts D183 and the α -carboxylate-S137 interaction becomes water-mediated (D183 and S137 in rat liver arginase number are equivalent to E188 and S129 in *H. pylori* arginase) [41]. The binding mode of the sulfonamide inhibitor may mimic the binding of the tetrahedral intermediate by displacing the metal-bridging hydroxide ion in the native enzyme and it is flanking transition states in catalysis. Our finding confirm the previous reports that the introduction of sulfonamide group increases polarity of a molecule and the hydrogen-bond donor properties as a sulfonamide N–H is more acidic ($\text{pK}_a = 11\text{--}12$) [19]. Furthermore, due to geometry of sulfur atom the sulfonamide bond shows structural similarity to the tetrahedral transition state present as an intermediate in the enzymatic hydrolysis of an amide bond thus making these compounds candidates in the development of lead compounds.

Table 2The interaction energy (kcal mol⁻¹) of *H. pylori* arginase and inhibitors obtained from the molecular dockings.

Substrate/inhibitor	Structure	Binding energy (kcal mol ⁻¹)	Docked energy (kcal mol ⁻¹)	Inter molecular energy (kcal mol ⁻¹)	Torsional energy (kcal mol ⁻¹)	Internal energy (kcal mol ⁻¹)	IC ₅₀ ^a
Arginine		-7.79	-9.42	-9.35	1.56	-0.07	ND ^b
NOHA		-6.86	-9.28	-9.04	2.18	-0.24	20
Nor-NOHA		-8.56	-10.71	-10.43	1.87	-0.28	0.8
Dinor-NOHA		-6.85	-8.2	-8.1	1.25	-0.1	40
Descarboxy-nor-NOHA		-6.18	-7.43	-8.42	1.25	0	300
(s)-2-Amino-7,7-dihydroxyheptanoic acid		-6.88	-8.75	-8.65	1.87	-0.1	ND
(s)-2-(Sulfonamidoethyl)-D-Cysteine		-6.62	-8.79	-8.48	1.87	-	ND
L-Valine		-6	-6.53	-6.62	0.62	0.09	650

^a IC₅₀ (μM) was derived from Christianson and colleague's study [41].^b ND, means not defined.

Thus, the above mentioned results not only corroborate the accuracy of the applied molecular modeling methods, but also confirm the 3D structure predicted of the *H. pylori* arginase and points to validity of this predicted geometry.

4. Conclusion

The aim of this study was to build a structural model of arginase in order to provide a reliable model with which to design new inhibitors and to investigate the role of some conserved residues between the *B. caldovelox* and *H. pylori* arginase and also modeling the binding mode of arginase with substrate and some of its inhibitor. After model construction, structural refinement by 30 ns

MD simulation and energy minimizations improved the general structure of the model and stable RMSD and total energy throughout the simulation were also obtained then various model evaluation methods indicated that the modeled structure has considerably good geometry. Despite the low sequence identity between the *H. pylori* and the *B. caldovelox* arginase which was used as a template, the model does have acceptable profiles for the three programs of structural analysis in all sequences except the region around residues 150–166 which is related to the region appeared as an insertion in the *H. pylori* that was not found in the alignment with other sequences. As in 1CEV, the catalytic site of *H. pylori* arginase is composed of residues from one subunit. Comparison between the modeled structures and crystal structure

of 1CEV showed similarity in active site topology and overall folding of enzymes. The docking studies of this homology model were also in consistence with the report by Mendz et al. [6], which shows that there are seven important conserved residues including: four aspartic acids; D116, D120, D234, D236 and three histidines; H91, H118 and H133. Those are of special importance for catalysis and stability of binuclear metal center of arginase and have significant role in binding and catalytic activity in the active site. Further docking analysis indicated that the binding model of highly active compounds was similar to that in the crystal structure of enzyme–ligand complexes. The docking energy were identified and shown to corroborate the experimental results. In the absence of the experimental structure of the *H. pylori* arginase we hope that our model will be useful to design new inhibitors that could bind to the proposed critical sites. Work in this context is in progress.

Acknowledgements

The authors would like to thanks Research council of Tehran University of Medical Sciences and Iranian National Science Foundation (INSF) for providing the financial support for this work.

References

- [1] B.J. Marshall, Unidentified curved bacilli on gastric epithelium in active chronic gastritis, *Lancet* 1 (1983) 1273–1275.
- [2] M.J. Blaser, *Helicobacter pylori* and the pathogenesis of gastroduodenal inflammation, *J. Infect. Dis.* 161 (1990) 626–633.
- [3] B.J. Marshall, D.B. McGechie, P.A. Rogers, R.J. Glancy, Pyloric campylobacter infection and gastroduodenal disease, *Med. J. Aust.* 142 (1985) 439–444.
- [4] G.E. Buck, W.K. Gourley, W.K. Lee, J.M. Subramanyan, Relation of *Campylobacter pyloridis* to gastritis and peptic ulcers, *J. Infect. Dis.* 153 (1986) 664–669.
- [5] J. Parsonnet, G.D. Friedman, D.P. Vandersteen, Y. Chang, J.H. Vogelstein, N. Orenreich, R.K. Sibley, *Helicobacter pylori* infection and the risk of gastric carcinoma, *N. Engl. J. Med.* 325 (1991) 1127–1131.
- [6] G.L. Mendz, E.M. Holmes, R.L. Ferrero, In situ characterization of *Helicobacter pylori* arginase, *Biochim. Biophys. Acta* 1338 (1998) 465–477.
- [7] G.L. Mendz, S.L. Hazell, The urea cycle of *Helicobacter pylori*, *Microbiology* 142 (1996) 2959–2967.
- [8] B.J. Marshall, L.J. Barrett, C. Prakash, R.W. McCallum, R.L. Guerrant, Urea protects *Helicobacter (Campylobacter) pylori* from the bactericidal effect of acid, *Gastroenterology* 99 (1990) 697–702.
- [9] J.E. Sjöström, H. Larsson, Factors affecting growth and antibiotic susceptibility of *Helicobacter pylori*: effect of pH and urea on the survival of a wild-type strain and a urease-deficient mutant, *J. Med. Microbiol.* 44 (1996) 425–433.
- [10] D.W. Christianson, J.D. Cox, Catalysis by metal-activated hydroxide in zinc and manganese metalloenzyme, *Annu. Rev. Biochem.* 68 (1999) 33–57.
- [11] W. Durante, F.K. Johnson, R.A. Johnson, Arginase: a critical regulator of nitric oxide synthesis and vascular function, *Clin. Exp. Pharmacol. Physiol.* 34 (2007) 906–911.
- [12] J.F.O. Tomb, O. White, A.R. Kerlavage, R.A. Clayton, G.G. Sutton, R.D. Fleischmann, K.A. Ketchum, H.P. Klenk, S. Gill, B.A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E.F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H.G. Khalak, A. Glodek, K. McKenney, L.M. Fitzgerald, N. Lee, M.D. Adams, E.K. Hickey, D.E. Berg, J.D. Gocayne, T.R. Utterback, J.D. Peterson, J.M. Kelley, M.D. Cotton, J.M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W.S. Hayes, M. Borodovsky, P.D. Karp, H.O. Smith, C.M. Fraser, J.C. Venter, The complete genome sequence of the gastric pathogen *Helicobacter pylori*, *Nature* 388 (1997) 539–547.
- [13] S.M. Morris, D. Bhamidipati, D. Kepka-Lenhart, Human type II arginase: sequence analysis and tissue-specific expression, *Gene* 193 (1997) 157–161.
- [14] C.A. Ouzounis, N.C. Kyrpides, On the evolution of arginases and related enzymes, *J. Mol. Evol.* 39 (1994) 101–104.
- [15] K. Ginalski, Comparative modeling for protein structure prediction, *Curr. Opin. Struct. Biol.* 16 (2006) 172–177.
- [16] M.C. Bewley, P.D. Jeffrey, M.L. Patchett, Z.F. Kanyo, E.N. Baker, Crystal structures of *Bacillus caldovelox* arginase in complex with substrate and inhibitors reveal new insights into activation, inhibition and catalysis in the arginase superfamily, *Structure* 7 (1999) 435–448.
- [17] J. Custot, C. Moali, M. Brollo, J.L. Boucher, M. Delaforge, D. Mansuy, J.P. Tenu, J.L. Zimmermann, The new α -amino acid *N*-omega-hydroxy-nor-L-arginine: a high-affinity inhibitor of arginase well-adapted to bind to its manganese cluster, *J. Am. Chem. Soc.* 119 (1997) 4086–4087.
- [18] C. Moali, M. Brollo, J. Custot, M.-A. Sari, J.-L. Boucher, D.J. Stuehr, D. Mansuy, Recognition of α -amino acids bearing various C=NOH functions by nitric oxide synthase and arginase involves very different structural determinants, *Biochemistry* 39 (2000) 8208–8218.
- [19] E. Cama, H. Shin, D.W. Christianson, Design of amino acid sulfonamides as transition-state analogue inhibitors of arginase, *J. Am. Chem. Soc.* 125 (2003) 13052–13057.
- [20] N. Carvajal, S.D. Cederbaum, Kinetics of inhibition of rat liver and kidney arginases by proline and branched-chain amino acids, *Biochim. Biophys. Acta* 870 (1986) 181–184.
- [21] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic Local Alignment Search Tool, *J. Mol. Biol.* 215 (1990) 403–410.
- [22] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucl. Acids Res.* 22 (1994) 4673–4680.
- [23] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, A. Sali, Comparative protein structure modeling using Modeller, *Curr. 2007 Chapter 2*, Unit 29.
- [24] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- [25] E. Lindhal, B. Hess, D. Van der Sipel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, *J. Mol. Mod.* 7 (2001) 306–317.
- [26] D.C. Bas, D.M. Rogers, J.H. Jensen, PROPKA, very fast prediction and rationalization of pK_a values for protein–ligand complexes, *Proteins* 73 (2008) 765–783.
- [27] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. Dinola, J.R. Haak, Molecular dynamic with coupling to an external bath, *J. Chem. Phys.* 81 (1984) 3684–3690.
- [28] D. York, W. Yang, The fast Fourier Poisson method for calculating EWALD sums, *J. Chem. Phys.* 101 (1994) 3298–3300.
- [29] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three dimensional profiles, *Nature* 356 (1992) 83–85.
- [30] M. Wiederstein, J.M. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucl. Acids Res.* 35 (2007) 407–410.
- [31] R. Maiti, G.H. Van Domselaar, H. Zhang, D.S. Wishart, SuperPose: a simple server for sophisticated structural superposition, *Nucl. Acids Res.* 32 (2004) 590–594.
- [32] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.* 19 (1998) 1639–1662.
- [33] C. Hetényi, D. van der Spoel, Efficient docking of peptides to proteins without prior knowledge of the binding site, *Protein Sci.* 11 (2002) 1729–1737.
- [34] A.R. Leach, *Molecular modelling, principles and applications*, second edition, 2004, pp 390–392.
- [35] J. Perozich, J. Hempel, S. Morris, Roles of conserved residues in the arginase family, *Biochim. Biophys. Acta* 1382 (1998) 23–37.
- [36] J. Hovey, E. Watson, M. Langford, E. Hildebrandt, S. Bathala, J. Bolland, D. Spadafora, G. Mendz, D. McGee, Genetic microheterogeneity and phenotypic variation of *Helicobacter pylori* arginase in clinical isolates, *BMC Microbiol.* 7 (2007) 26–41.
- [37] D.W. Christianson, R.S. Alexander, Carboxylate–histidine–zinc interactions in protein structure and function, *J. Am. Chem. Soc.* 111 (1989) 6412–6419.
- [38] Z.F. Kanyo, L.R. Scolnick, D.E. Ash, D.W. Christianson, Structure of a unique binuclear manganese cluster in arginase, *Nature* 383 (1996) 554–557.
- [39] A. Sekowska, A. Danchin, J.L. Risler, Phylogeny of related functions: the case of polyamine biosynthetic enzymes, *Microbiology* 146 (2000) 1815–1828.
- [40] D.J. Bacon, W.F. Anderson, A fast algorithm for rendering space-filling molecule pictures, *J. Mol. Graph.* 6 (1988) 219–220.
- [41] E. Cama, S. Pethe, J.L. Boucher, S. Han, F.A. Emig, D.E. Ash, R.E. Viola, D. Mansuy, D.W. Christianson, Inhibitor coordination interactions in the binuclear manganese cluster of arginase, *Biochemistry* 43 (2004) 8987–8999.