STUDIA UNIVERSITATIS BABEŞ-BOLYAI, CHEMIA, XLV, 2, TOM I, 2010

Dedicated to the memory of Prof. dr. Ioan Silaghi-Dumitrescu marking 60 years from his birth

ZINC-CONTAINING ACTIVE SITE IN A PARTIALLY MODIFIED 1GKM CRYSTAL STRUCTURE OF HISTIDINE AMMONIA-LYASE: A COMPUTATIONAL INVESTIGATION

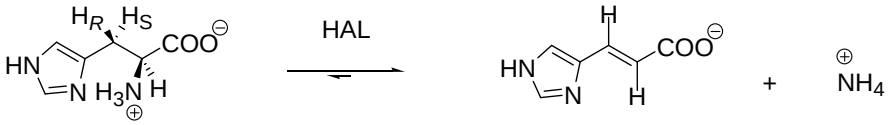
AMALIA-LAURA SEFF[a], SAROLTA PILBÁK[b], IOAN SILAGHI-DUMITRESCU[a], LÁSZLÓ POPPE[b]

ABSTRACT. The histidine ammonia-lyase (HAL) biocatalyst has an important role in the formation of (*E*)-urocanic acid from L-histidine eliminating ammonia. In order to investigate the active center of HAL, we have prepared "*in silico*" a HAL model with a compact active center on the basis of the recently appeared phenylalanine ammonia-lyase (PAL) structure (PDB code: 3CZO). Furthermore, to have a better view on the orientation of a key reaction intermediate within the active site, systematic conformational analysis has been performed. Furthermore, the possible positions of the L-histidine/(*E*)-urocanic acid were evaluated by optimization and docking.

Keywords: *histidine ammonia-lyase, homology modeling, conformational analysis, docking, Zn^{2+}*

INTRODUCTION

The degradation pathway of histidine in various organisms proceeds in a different way than for most of the other amino acids [1]. The histidine ammonia-lyase (HAL, EC 4.3.1.3) catalyzes the first step of the degradation with ammonia elimination from the L-histidine substrate forming the product (*E*)-urocanic acid (Scheme 1).



L-histidine (*E*)-urocanic acid

Scheme 1

[a] *Faculty of Chemistry and Chemical Engineering, Babeş-Bolyai University, Kogălniceanu Str. No. 1, RO-400084 Cluj-Napoca, Romania, samalia@chem.ubbcluj.ro*
[b] *Department of Organic Chemistry and Technology, Budapest University of Technology and Economics, H-1111 Budapest, Gellért tér 4, Hungary, poppe@mail.bme.hu*

The aforementioned product was identified as a component of human sweat [2] and a condition named histidinemia appears in case of the lack of HAL in humans [3]. Histidine ammonia-lyase together with the phenylalanine ammonia-lyase (PAL, EC 4.3.1.24) and tyrosine ammonia-lyase (TAL, EC 4.3.1.23) are a part of the family of ammonia-lyases. PAL and TAL as members of this family catalyze the elimination of ammonia from the corresponding L-amino acid forming unsaturated products as (*E*)-cinnamic acid and (*E*)-coumaric acid. These products have important biological roles from the amino acid catabolism to the formation of intermediates in different biosyntheses. HAL, PAL and TAL catalyze the ammonia elimination by the aid of the 5-methylene-3,5-dihydroimidazol-4-one (MIO) prosthetic group [4].

It was observed that cations, like Zn^{2+}, Cd^{2+} or Mn^{2+} can increase the activity of the HAL enzyme [5]. After elucidation of the X-ray structure of HAL [6], the importance of several amino acid residues at the active center (Tyr53, His83, Asn195, Gln277, Tyr280, Arg283, Phe329 and Glu414) for catalysis and substrate binding was evaluated by site-directed mutagenesis and it was supposed that Zn^{2+} may play a role in catalysis [7].

The present work describes the "*in silico*" preparation of the partially modified HAL structure (1GKM*mod*) from the crystal structure (1GKM) [8] of *Pseudomonas putida* HAL (*Pp*HAL), including modeling of a closed Tyr53-loop region which was necessary to obtain a closed active site. A systematic conformational analysis of the substrate covalently bound to the MIO prosthetic group (*N*-MIO intermediate) taking into account the presence of a Zn^{2+} cation and ligand docking results are also presented.

RESULTS AND DISCUSSION

Selection and modeling of the HAL crystal structure

According to the study with active site mutants of HAL [7], Tyr53, Glu414 and Tyr280 residues may have an important role in the abstraction of the pro-(S) β-proton from the L-histidine substrate. A comparison of the structure of *Pp*HAL to other ammonia-lyases showed that the loop with the essential Tyr53 adopts a partially open conformation (Figure 1).

The X-ray structure of plant PAL from *Petroselinum crispum* (PDB code: 1W27) determined at 1.7 Å resolution [9] has an opened, solvent accessible active site (Figure 1A). The bacterial *Pp*HAL (PDB code: 1GKM) structure reveals a partially opened, solvent accessible active site (Figure 1B). All the six crystal structures determined for HAL so far contain the catalytically essential Tyr53 in a partially open loop conformation. This may be the reason why these structures could not retain subtrate or product related ligands. Therefore, none of them could be used for theoretical investigation of the compact active center. Only the 1GKM crystal structure of HAL (inhibited with L-cysteine) was determined in the presence of an inhibitor in the active center.

The recently published ammonia-lyase crystal structure (PDB code: 3CZO) [10], determined for Anabaena variabilis PAL (AvPAL), contains the most compact active center in which the essential Tyr78 and the MIO prosthetic group are deeply buried and not solvent accessible (Figure 1C).

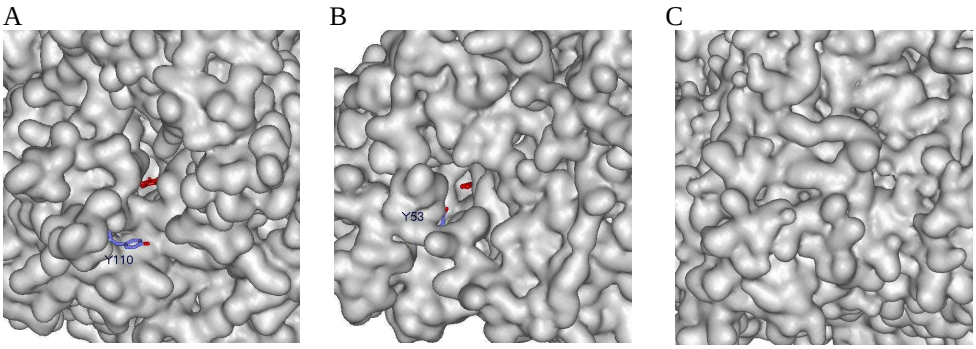


Figure 1. Comparison of the substrate entrance chanel towards MIO group (red) in several ammonia-lyase structures. Representation of molecular surface of (A) *Pc*PAL (PDB code 1W27), (B) *Pp*HAL (PDB code 1GKM) and (C) *Av*PAL (PDB code 3CZO) crystal structures. The analogous Tyr110 (A) and Tyr53 (B) residues are seen as stick models in the partially opened ammonia-lyase structures.

Our motivation to select the 1GKM HAL structure for further calculations was the presence of an inhibitor and the unique orientation of the side chain of Met382 amino acid within the active site.

Modification of the essential Tyr-containing loop region (39-80 amino acid sequence) of 1GKM HAL was based on the 3CZO PAL structure using homology modeling (Figure 2). The new (1GKM$_{mod}$) model contains a tightly closed active site and thus provides a well defined enzymic ambience for further theoretical investigation and calculations.

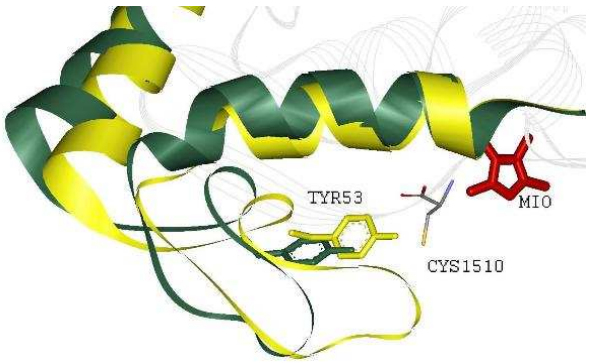


Figure 2. Overlay of the important Tyr53 loop regions of *Pp*HAL (1GKM, green chain; including the L-cysteine inhibitor in CPK color) and of the partially modified 1GKM$_{mod}$ (yellow chain) structures.

Conformational analysis of the N-MIO intermediate

The arrangement of L-cysteine inhibitor within the active center of 1GKM crystal structure provided the opportunity to replace the inhibitor with the natural substrate (L-histidine). In the *N*-MIO intermediate of the HAL reaction, the L-histidine is covalently bound to the MIO prosthetic group trough its amino group. The carboxyl group of the histidine substrate may point towards the Arg283 amino acid residue and the imidazole ring of the substrate may be in the hydrophobic part of the active site close to the His83 residue. In order to get the geometries of the possible orientation of the intermediate within the active site, the systematic conformational analysis search (CS) method has been used. The active center of the HAL structure was proposed to contain a Zn^{2+} ion [7]. Concerning on the Zn-binding site, the S atom of Met382 could be one of the coordinating ligand of the Zn^{2+} ion together with the imidazole group of His83 and substrate. The W544 water molecule was moved as the fourth ligand of Zn^{2+}.

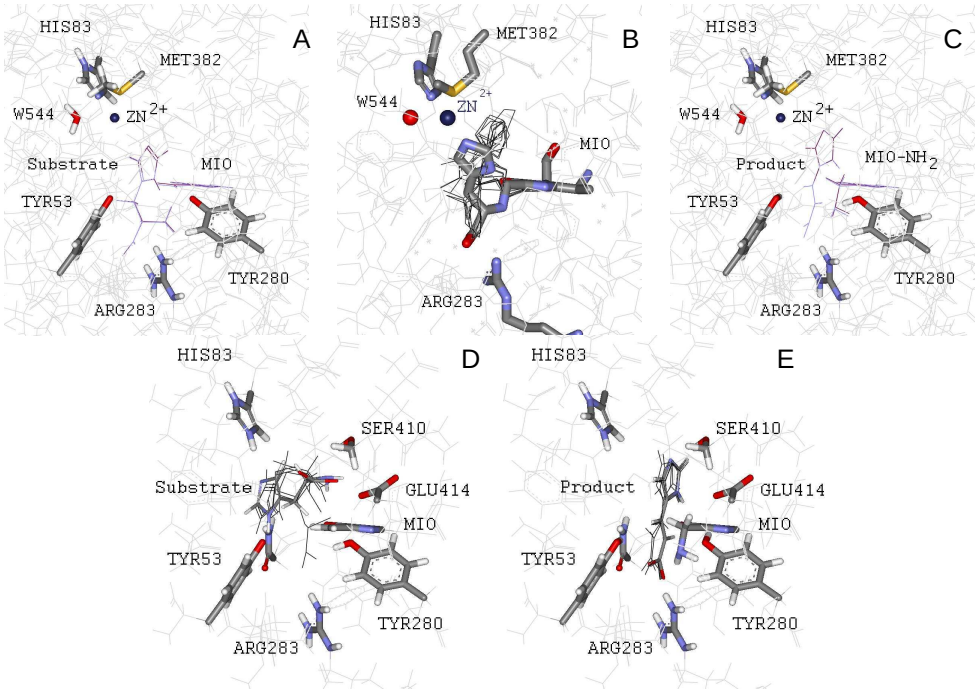


Figure 3. Optimized arrangements of substrate (A) and product (C), conformational analysis of the covalent intermediate (B), and docking results for the substrate (D) and product (E) in the modified HAL structure.

For the covalent intermediate model (Figure 3B), the best ligand energy conformation was selected and highlighted as stick model (imidazole group points towards His83 and carboxylate is in the close vicinity of Arg283) which was also in correspondence with the overall arrangement of the ligands found in the experimental structures of *Rhodobacter sphaeroides* TAL. The Zn^{2+} coordination set is formed by the imidazole group of His83 and of the substrate, the Met382 and a water molecule. The geometrical characterization of the optimized substrate, covalent intermediate and product structures is presented in the Table 1.

Table 1. Geometrical parameters resulted after optimization of the substrate/product structure and after conformational search of the covalent intermediate structure.

	Distances (Å)			
	d(Zn-$N3_{H83}$)	d(Zn-$N3_{Lig}$)	d(Zn-O_{W544})	d(Zn-S_{M382})
substrate	1.91	2.90	2.91	3.76
intermediate	1.91	2.67	2.91	3.76
product	1.91	3.05	2.91	3.76
	d($H_{O\text{-}Y53}$-$C\beta_{Lig}$)	d(βH_{Lig}-O_{Y53})	d(O_{Lig}-H_{R283})	d($H_{N1\text{-}Lig}$-O_{E414})
subst./prod.	2.21	2.20	1.85/2.01	3.00/3.28
	Angles(º)			
	$N3_{H83}$-Zn-$N3_{Lig}$	$N3_{Lig}$-Zn-O_{W544}	O_{W544}-Zn-S_{M382}	S_{M382}-Zn-$N3_{H83}$
substrate	165.5	113.7	47.1	84.8
intermediate	164.0	86.8	47.1	66.6
product	147.2	117.6	47.1	70.4

Geometry optimization and ligand docking of the substrate/product

Geometry optimization of the substrate/product ligands within the active site of $1GKM_{mod}$ was necessary because none of the crystal structures determined for HAL contain the natural L-histidine substrate or the (*E*)-urocanic acid product of the reaction.

For a good evaluation of the substrate/product position within the active site, molecular mechanics (MM+) optimization (Figure 3A and 3C) of the ligands derived from the best covalent *N*-MIO intermediate using HyperChem program [11] and ligand docking (Figure 3D and 3E) within the active center of $1GKM_{mod}$ structure using AutoDock software [12, 13] were carried out.

CONCLUSIONS

Taking into account that partial modification of the 1W27 *Pc*PAL crystal structure by homology modeling [14, 15, 16] was already applied for computational investigation within the active site of the reaction mechanism, it can be considered that our novel HAL model (1GKM$_{mod}$) is eligible for further computational investigation of the possible geometry orientation of the L-histidine substrate/*N*-MIO intermediate/(*E*)-urocanic acid product.

Although the ligand docking is a computational tool which is dedicated to the evaluation of the possible positions of the ligand within the active site using different algorithms, the results obtained at the L-histidine and (*E*)-urocanate docking within the 1GKM$_{mod}$ active center revealed only for the (*E*)-urocanate product the following interactions: the carboxylate moiety establishes an interaction with the Arg283 and the imidazole ring points towards the His83 amino acid residue trough its N_3 atom.

EXPERIMENTAL SECTION

Selection of the HAL crystal structure for modification and modeling

There are six crystal structures determined for *Pseudomonas putida* histidine ammonia-lyase enzyme (Brookhaven Protein Data Bank codes: 1B8F, 1EB4, 1GK2, 1GK3, 1GKJ, 1GKM). Among the available structures the 1GKM structure was selected for computational investigation of HAL. This is the only structure which contains a ligand (L-cysteine) inside the active center and revealed a unique orientation of the side chain of Met382 amino acid within the active site compared to the other structures determined for HAL. The distance between the N atom of the amino moiety in L-cysteine and the C atom of the methylene group in MIO was 1.31 Å.

The 3CZO *Av*PAL crystal structure has the inner Tyr loop region (containing the essential Tyr53 amino acid) in a fully closed form. The 39-80 amino acid sequence of 1GKM HAL, which contains the important Tyr53 amino acid in a partially open loop conformation, was replaced by the corresponding homology model sequence of the 1GKM HAL based on 3CZO PAL. The resulted 1GKM$_{mod}$ model structure seemed to be eligible for calculations within the enzymic ambience taking into account the new position of the Tyr53 amino acid in comparison with the one from the crystal structure of *Pp*HAL within the active site (Figure 2).

Conformational analysis of the N-MIO intermediate

The initial ligand structures for the CS were built from the L-cysteine inhibitor (Cys1510) of the 1GKM crystal structure of *Pp*HAL. Therefore, the L-cysteine of the 1GKM was replaced by L-histidine in the 1GKM$_{mod}$ HAL structure. In the resulted starting conformation the carboxyl group of the MIO-bound L-histidine pointed towards the Arg283 residue and the imidazole

ring of the substrate occupied a relative hydrophobic part of the active site close to the His83 residue. From the resulted 1GKM$_{mod}$ structure a region defined by a sphere of 15 Å radius around the C atom of the original L-cysteine (Cys1510) and a sphere of 12 Å radius around the O atom of a highly conserved water (W544 – close to the place where we propose the presence of the Zn^{2+} ion) was cut off for CS calculations as a closed active site model of HAL (including 146 amino acids and 43 water molecules). Next, by a HyperChem [11] standard procedure, hydrogen atoms were added to the amino acid moieties. In the starting structure, the *C*- and *N*-termini at cuttings were completed to neutral aldehyde and amino parts. The resulted overall structure and the hydrogen additions were verified and corrected [17]. From this verified raw *Pp*HAL active site model (partially modified 1GKM crystal structure including L-histidine), only one initial active site construct involving the substrate covalently bound to MIO trough the amino moiety (*N*-MIO model) was built. In this active site model a Zn^{2+} ion was included at a 1.9 Å distance from the N3 atom of His83 amino acid residue (as found in many Zn-containing protein structures). During CS on the covalently bound histidine and the imidazolone ring of the MIO in rigid enzymic environment (including 41 water molecules), 3 of the ligand torsion angles were varied for the *N*-MIO model. CS's were performed by HyperChem implemented CS module [11] using default settings (MM+ forcefield; gradient: 0.1 kcal/mol; Polak-Ribiere method; limits: 300 iterations, 150 optimizations, 15 conformations; test options: "skip if atoms are closer than 0.3 Å").

For the covalent intermediate model, the best ligand energy conformation was selected which was also in agreement with the overall arrangement (imidazole group points towards His83 and carboxylate is in the close vicinity of Arg283) found in the experimental inhibited structures of *Rhodobacter sphaeroides Rs*TAL.

Several *N*-MIO active site models were constructed by replacing the substrate-MIO part of the initial *Pp*HAL (1GKM$_{mod}$)/ L-histidine active site construct with the ligand arrangements resulted from the CS.

Geometry optimization and docking of L-histidine/(E)-urocanic acid in the active site model of PpHAL

In order to construct the L-histidine/(*E*)-urocanate ligand-containing starting structures, six different conformers resulted from the CS were used. The covalent bond (Figure 4) between the methylene group of the MIO and the amino moiety was cut in case of the substrate (indicated by the green line). The bond between the amino moiety and the carbon atom of the ligand was cut in case of the product (indicated by the blue line) and the structure was completed with the necessary olefinic bond to get (*E*)-urocanate.

Next, the 6 × 2 L-histidine/(*E*)-urocanate ligand structures were optimized within rigid enzymic environment, by the HyperChem implemented *Geometry optimization* using the default settings (MM+ forcefield; gradient: 0.1 kcal/mol; Polak-Ribiere method). From the resulted set of models, the two lowest energy substrate and product structures (ones obtained from the c12 and c14 optimizations) were used for further optimizations in which the essential parts of the active site (including the MIO-NH_2, the aromatic part of the His83, Tyr280, Tyr53, the $–CH_2$-CO-NH_2 part of Asn195, Asn313, Gln277, Gln413, the $–CH_2$-COO^- part of the Glu414, the -CH_2-OH part of Ser410 and the Zn^{2+} ion) were included.

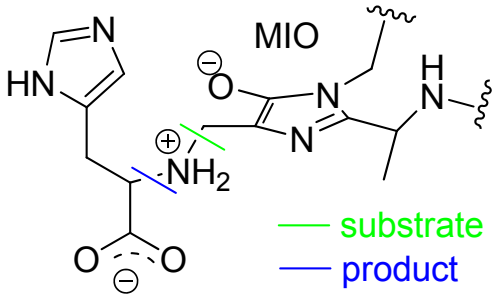


Figure 4. The bond cuts of the *N*-MIO intermediate to construct the L-histidine substrate (green) and (E)-urocanate product (blue) containing HAL models

The starting structures for the docking calculations were constructed from the active site model of *Pp*HAL 1GKM$_{mod}$ structure used for CS calculations. Only 29 amino acid residues (including the essential amino acid residues like His83, Arg283, Met382, Tyr53, Tyr280, Asn195 and Glu414, determined by mutagenesis) and the MIO prosthetic electrophilic group were used as environment for the substrate/product docking. We used separate initial active site models for the substrate (4 torsion angle were selected from the substrate using flexible-rigid docking) and for the product. For the docking process of L-histidine/(*E*)-urocanic acid in the 1GKM$_{mod}$ active site models, the AutoDock software [12, 13] was used. The Grid Parameter File was established with the default settings (Total Grid Pts per map: 77326; number of points 40 for x- and y-dimensions, 45 for z-dimension; spacing: 0.375 Å; center on ligand). The docking process of the L-histidine/(*E*)-urocanic acid in 1GKM$_{mod}$ active site model was performed by the AutoDockTools implemented Docking module using default settings (Lamarckian Genetic Algorithm also known as a Genetic Algorithm-Local Search: 10 runs; population size: 150; maximum number of energy evaluations: short setting).

For the product-containing $1GKM_{mod}$ active site model, a set of the best docking energy poses were selected (Figure 3E), which were also in correspondence to the overall arrangement found for (*E*)-coumarate in wild-type *Rhodobacter sphaeroides bacterial* TAL (PDB code: 2O7B).

ACKNOWLEDGMENTS

ALS thanks the financial support of Romanian National University Research Council (CNCSIS, PNII-TD-400). *PL* thanks OTKA T048854 for financial support.

Supporting information available:

The coordinates of the $1GKM_{mod}$ structure.

REFERENCES

1. L. Poppe, J. Rétey, *Angewandte Chemie, International Edition English*, **2005**, *44*, 3668.
2. H. Morrison, C. Bernasconi, G. Pandey, *Photochemistry and Photobiology*, **1984**, *40*, 549.
3. R.G. Taylor, H.L. Levy, R.R. McInnes, *Molecular Biology and Medicine*, **1991**, *8*, 101.
4. L. Poppe, *Current Opinion in Chemical Biology*, **2001**, *5*, 512.
5. C.B. Klee, *The Journal of Biological Chemistry*, **1972**, *247*, 1398.
6. T.F. Schwede, J. Rétey, G.E. Schulz, *Biochemistry*, **1999**, *38*, 5355.
7. D. Röther, L. Poppe, S. Viergutz, B. Langer, J. Rétey, *European Journal of Biochemistry*, **2001**, *268*, 6011.
8. M. Baedeker, G.E. Schulz, *European Journal of Biochemistry*, **2002**, *269*, 1790.
9. H. Ritter, G.E. Schulz, *Plant Cell*, **2004**, *16*, 3426.
10. L. Wang, A. Gamez, et al., *Journal of Molecular Biology*, **2008**, *380,* 623.
11. HyperChem version 7.0 (Hypercube, Inc. http://www.hyper.com/).
12. G.M. Morris, D.S. Goodsell, et al., *Journal of Computational Chemistry*, **1998**, *19*, 1639.
13. R. Huey, G.M. Morris, et al., *Journal of Computational Chemistry*, **2007**, *28*, 1145.
14. S. Pilbák, A. Tomin, J. Rétey, L. Poppe, *The FEBS Journal*, **2006**, *273*, 1004.
15. A.L. Seff, S. Pilbák, L. Poppe, *Studia Universitatis Babeş-Bolyai, Chemia* LIII, 2**008**, *2*, 67.
16. S. Bartsch, U.T. Bornscheuer, *Angewandte Chemie, International Edition English,* **2009**, *48*, 3362.
17. Molfunction, Institute of Molecular Function (http://www.molfunction.com/).