Research Article [Araştırma Makalesi]

Implication of molecular conservation on computational designing of haloarchaean urease with novel functional diversity

[Geliştirilmiş fonksiyonel çesitlilik açısından haloarkea üreaz'ının hesaplamalıtasarımında moleküler korumanın etkisi*]

Paulchamy Chellapandi[1],

Jayachandrabal Balachandramohan[1]

[1]Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli-620024, Tamil Nadu, India

ABSTRACT

Objective: The objective was to design an enzyme construct with diverse function from urease sequences of haloarchaean, *Haloarcula marismortui* ATCC 43049 based on its conserved domain consisting metal-binding region and active sites.

Methods: Complete urease sequences of haloarchaea were retrieved from National Center for Biotechnology Information and then homology models generated, and validated. The best protein models were selected for docking with respective substrates using Ligand Fit program. The lowest energetic conformers were generated from these protein models by molecular dynamics methods. Urease construct-substrate complex was chosen based on the mode of catalysis, types of molecular interactions, and binding energy.

Results: The resulted construct has a monomeric structure consisting of 3 helixes and 6 turns with 97 amino acids in length. The side chains of Asp49, Gly50 and Gln51 were predicted as functional residues in this construct. Urease construct was predicted to show catalytic function as similar to aliphatic nitrile hydradase and acrylamide hydro-lyase. Binding affinity of construct was more significant, which was better than to native urease. Urease construct was showed high binding affinity with semicarbazide and acrylamide wherein it has formed favorable hydrogen bonds.

Conclusion: Substrate-binding region and active sites in the conserved domain of haloarchaean ureases are evolutionarily conserved at sequence as well as structural level. Substrate docking study supports the strong molecular interactions between construct and relative substrates. Thus, the present approach provides an insight to design urease construct with diverged catalytic function.

Key Words: Molecular docking, urease, nitrilase, molecular evolution, enzyme design, haloarchaea

Conflict of interest: Authors have no conflict of interest.

Yazışma Adresi
[Correspondence Address]

Paulchamy Chellapandi

Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli-620024, Tamil Nadu, India
Tel. +91-431-2407071
Fax. +91-431-2407045
Email: pchellapandi@gmail.com

ÖZET

Amaç: Bu çalışmada, *Haloarcula marismortui* ATCC 43049'in metal bağlayan bölge ve aktif bölge gibi korunmuş domainleri esas alınarak haloarchaeanın üreaz dizilimine ters fonksiyonlu enzim yapısının tasarlanması amaçlandı.

Yöntem: Haloarchaea ait üreaz dizisi, Ulusal Biyoteknoloji Bilgi Merkezinden alındı ve homolog modeller geliştirilerek doğrulama yapıldı. Docking için en uygun protein modelleri Ligand Fit programı kullanılarak ilgili substratlar ile seçildi. Bu protein modellerinden en düşük enerjili olan uyumlu modeller moleküler dinamik metodlar ile oluşturuldu. Üreaz yapı-substrat kompleks modeli enerji bağlama kapasitesi, moleküler etkileşim tipleri ve kataliz tipi esas alınarak seçildi.

Bulgular: Meydana getirilen yapı 97 amino asit uzunluğunda 3 heliks ve 6 dönüş içeren monomer özelliğindedir. Bu yapıda Asp49, Gly50 ve Gln51 yan zincirleri fonksiyonel kalıntılar olarak öngörüldü. Üreaz yapısının ise alifatik nitril hidrataz ve akrilamid hidroliyaza benzer katalitik fonksiyonları olduğu düşünüldü. Oluşturulan üreazın bağlanma eğilimi esas üreazdan daha iyi olarak anlamlıydı. Ayrıca üreaz semikarbazid ve akrilamid ile yüksek bağlanma eğilimi gösterdi ve uygun hidrojen bağları oluşturdu.

Sonuç: Haloarchaean üreaza ait korunmuş domainde bulunan substrat bağlayıcı bölge ve aktif bölge yapısal olduğu kadar, dizilim düzeyinde de evrimsel olarak korunmuştur. Substrat docking çalışmaları oluşturulan yapı ve ilgili substrat arasında sağlam moleküler etkileşimlerin olduğunu desteklemektedir. Bu çalışma üreaz yapısının tasarlanmasına farklı katalitik fonksiyonlar üzerinden değişik bakış açıları sağlamaktadır.

Anahtar sözcükler: Moleküler docking; üreaz, nitrilaz; moleküler evrim; enzim tasarımı; haloarchaea

Çıkar çatışması: Yazarlar herhangi bir çıkar çatışması bildirmemektedirler.

Introduction

The maturation of enzyme technology is shown by the development of the theory concerning how enzymes function and how this is related to their primary structure through the formation and configuration of their three-dimension structure [1,2]. The design of artificial enzyme is based on the knowledge about the structure, architecture and functional properties of biological enzymes. It is well known that the enzymes contain a binding site and a catalytic site consisting of two or more catalytic amino acid groups [3,4]. Exploitation of the diverse reactivities of metal center cofactors presents a profitable strategy to introduce catalytic activity into proteins. Several different potential reactivities toward a single substrate often exhibit on metal centre [5,6]. Hence, computer-aided enzyme modeling has taken an important effort to design metalloenzymes so as to perform chemical reactivity with good catalytic efficiency in biotransformation processes.

Nickel is a key metal involved in many of the biochemical process in archaea, and urease (urea amidohydrolase; EC 3.5.15) is one of the nickel-dependent metalloenzymes in haloarchaea. Apart from urease, other archaeal nickel-dependent enzymes are more diverse in nature so that urease has taken more advantages for a rational enzyme designing [7,8]. Urease catalyzes the hydrolysis of urea to yield ammonia and carbamate, which spontaneously hydrolyzes to form carbon dioxide and a second molecule of ammonia [9]. It is composed of three subunits, encoded by the genes *ure*A, *ure*B, and *ure*C. The biosynthesis of a functional urease also requires the presence of four additional genes (*ure*DEFG) [10]. The gene *ure*E encodes a nickel carrier protein [11], while *ure*DFG encode a chaperone complex that keeps urease in a configuration competent to accept a nickel ion and also requires carbamylation for efficient nickel incorporation [12,13].

Unfortunately, naturally available enzymes are usually not optimally suited for industrial applications due to the less stability under process conditions, when applying them in biotransformation reactions in industry [14]. Though protein engineering technologies can be used to alter variety of enzyme properties simultaneously, the appropriate screening parameters such as mutant library construction and variant selection should be employed [2]. Hence, the successful designs of small (less than 75 residues) monomeric proteins [15], protein oligomers [16], and the redesign of natural proteins to confer novel functionalities [17] have been achieved by the development and use of computational methods for searching the sequence space associated with a particular target structure.

The generation of active biocatalysts from dramatically reduced amino acid alphabets provides a strong support for the idea that primordial enzymes are made from only a handful of building blocks [1,14,16]. The binding of a substrate close to functional groups in the enzyme causes catalysis by so-called proximity effects. The success of current protein design methods based largely on optimizing the molecular energy potentials suggested that the proposed natural design properties are not necessary conditions for producing well-folded and perhaps even functional artificial proteins [17, 18]. It is therefore possible to design similar biocatalysts from small molecule mimics of enzyme active sites by combining in a small molecule and evolutionary conservation of sequences. In this context, we have aimed to use computer-aided modeling of urease constructs with diverse substrate-specificity based on evolutionary conservation of urease sequences at nickel- and substrate-binding regions.

Materials and Methods

Evolutionary conservation analysis

Complete haloarchaean urease sequences were retrieved from GenPept of National Center for Biotechnology Information (NCBI). Multiple sequence alignment was carried out for selected sequences with complete deletion of gaps and correction in multiple substitutions using ClustalX 2.0 software [19]. The aligned sequences were iterated at each alignment step and manually inspected to delete the low scoring sequences. Homogeneous patterns among all sequences were searched by Neighbor joining (NJ) algorithm to construct a phylogenetic tree with 1000 bootstraps values using MEGA 4.0 software [20]. The NJ algorithm calculates distances (percent divergence) between all pairs of sequence from a multiple alignment and applies it to the distance matrix. Because NJ method only gives strictly dichotomous trees (never more than 2 sequences join at one time), a multifurcation (several sequences joining at the same part of the tree) cannot be exactly represented. Using conserved domain search tool [21], conserved domains architecture as well as metal-binding templates of query sequences was searched from NCBI-CDD (Conserved Domain Database) [22].

Molecular modeling and enzyme designing

PSI-BLAST tool with a default parameter was used to search suitable protein data bank (PDB) templates for structure modeling from the sequences [23]. ModWeb is an automatic comparative protein modeling server which was used to build three dimensional (3D) structures from query sequences [24]. It enables a thorough exploration of fold assignments, sequence–structure alignments and conformations, with the aim of finding the model with the best evaluation score. A representative model for each alignment is chosen by ranking based on the atomic distance-dependent statistical potential Discrete Optimized Protein Energy (DOPE). The fold of each model is evaluated using a composite model quality criterion that includes the coverage of the modeled sequence, sequence identity

implied by the sequence–structure alignment, the fraction of gaps in the alignment, the compactness of the model and various statistical potential Z-scores. Active site residues of selected models were predicted by ProFunc server, which helps to identify the likely biochemical function of a protein from its 3D structure [25]. Crystallographic protein structures whose catalytic domains are similar to metal- and substrate-binding sites were compared with the models. Amino acid residues exclude metal- and substrate-binding regions and active sites have been removed from modeled proteins through atomic coordinates. Amino acid residues corresponding to the selected atomic coordinates were further used to generate 3D homology modeling structure using Prime program in Maestro software package (Schrodinger Inc.). The resulting model was evaluated using Structural Analysis and Verification Server (SAVS) (http://nihserver.mbi.ucla.edu/SAVES/), and then superimposed on the corresponding PDB template with Dali pairwise comparison tool in DALITE server (http://ekhidna.biocenter.helsinki.fi/dali_lite/start).

Molecular dynamics simulation of enzyme constructs

Structural conformers of the models were generated by Discovery Studio software using CHARMM force field, and steepest descent as well as adopted basis Newton-Raphson algorithms. Distance constraint was between N-terminal to C-terminal and dihedral restraint was started from C to $C\alpha$ (Φ) of first amino acid residue and $C\alpha$ to N (ψ) of second amino acid residue until the last amino acid residue in a molecular dynamic ensemble. After molecular dynamic simulation, the energy conformer 1 (lowest one) was selected for computing binding energies of construct-substrate complexes.

Molecular docking studies

Urea-related substrate structures in MOL2 files were retrieved from KEGG database using SIMCOM software (http://www.genome.jp/tools/simcomp) and then converted to PDB format. AutoDock 4.0 software, implemented with Genetic algorithm and AMBER force field, was used to dock substrate into construct. Genetic algorithm is adaptive heuristic search premised on the evolutionary ideas of natural selection and its basic concept is designed to simulate processes in natural system necessary for evolution, specifically those that follow the principles first laid down by Charles Darwin of survival of the fittest. Binding site (cavity) of each construct was selected within an enegy grid and a flexible substrate prefered to dock into it using Ligand Fit program. Smart energy minimization algorithm was used to refine the orientation of the substrate in the receptor site after finding good docking models. The quality of docking models was evaluated by computing interaction distances, binding energy terms and inhibition constants of each construct-substrate complex.

Results

Analysis of molecular conservation

Eight nickel-dependent enzymes (coenzyme F_{420} reducing hydrogenase, F_{420} non-reducing hydrogenase, methyl-coenzyme M reductase, hydrogenase maturation protease, carbon monoxide dehydrogenase, rubredoxin, urease and acetyl-CoA decarbonylase/synthase) were entries available for archaea in NCBI database. Text mining of this study pointed out many urease sequences including alpha, beta and gamma subunits for haloarchaea among archaeal domain. Protein sequences (NCBI accession YP_134542, BAC84959 and Q75ZQ4) have shown a good structural identity with corresponding crystallographic structures, which was ranged from 58 to 60% (Table 1). Metal-binding domain of these sequences was existed at the position 5-85 amino acids corresponding to the PDB template 2FVH (A). Urease sequence of *Haloarcula marismortui* ATCC 43049 (accession YP_134542) was most likely suited for rational enzyme design because of it has the shortest amino acid length to cover metal-and substrate-binding sites. The actual length of selected region for modeling was 95 amino acid residues. Construct was predicted to show similarity (e-value 1.27e-38; bit score 153; CD length 96 amino acids) to Uraese_gamma subunit (CD00390), a nickel dependent metalloenzymes (Figure 1). Amino acids Asp49, Gly50 and Gln51 were predicted as active site (nest) residues that were similar to PDB template 2FVH (A). Conservation score of predicted functional residues was 2.092 (Table 2). Due to a low identity and modeling score, positions beyond active sites and substrate-binding regions, the rest of the modeled proteins have been neglected from this study.

Phylogenetic analysis

In phylogenetic tree, the sequences of urease from haloarchaea were formed three separate clades such as subunits of alpha, beta and gamma. The sequence (construct) of *H. marismortui* ATCC 43049 was typically clustered within halophilic archaea and then with *Metallosphaera sedula* DSM 5348 (Figure 2). Gamma ureases were shared their phylogenic resemblance with alpha and beta ureases of haloarchaea and showed their functional uniqueness. As the sequences of alpha and beta ureases were distantly related with gamma urease, a clade formed by them was not included in this phylogenetic tree.

Structural quality and accuracy of urease construct

The sequence of urease construct was highly similar to the PDB template 2FVH (A) wherein we calculated sequence identity 60.4%, e-value 7.23e-19 and total energy -4061.003 kJ/mol. When its homology model was superimposed with 2FVH (A) (urease, gamma subunit; 1.80Å), it was predicted to show 21.6 Z-score

Table 1. Homology modeling data for predicting 3D structure from urease sequences

NCBI Accession	No. of Amino acid	Template ID	Identity (%)	Target Position	MPQS*	Z-Dope**
Urease alpha-subunit						
YP_001190983	555	4ubpC	56	1-555	1.69	-0.75
Q18EB9	570	1a5lC	57	5-570	1.63	-0.5
Q75ZQ5	568	1a5lC	57	4-568	1.62	-0.38
BAC84958	568	1a5lC	57	4-568	1.62	-0.38
YP_134541	568	1a5lC	57	4-568	1.62	-0.38
Q3IRZ5	570	1a5lC	57	5-570	1.63	-0.42
Urease beta-subunit						
CAJ53713	126	1ejxB	57	6-104	1.51	-0.64
Q75ZQ6	138	4ubpB	58	5-108	1.49	-0.66
BAC84957	138	4ubpB	58	5-108	1.49	-0.66
YP_134540	138	4ubpB	58	5-108	1.49	-0.66
Q3IRZ6	132	1ejxB	63	9-109	1.52	-0.37
ABP95060	215	4ubpA	59	1-99	1.28	-1.36
YP_001190984	215	4ubpA	59	1-99	1.27	-1.28
Urease gamma-subunit						
YP_659289	108	4ubpA	58	1-99	1.77	-1.84
Q75ZQ4	128	2fvhA	60	2-99	1.66	-1.97
BAC84959	128	2fvhA	60	2-99	1.66	-1.97
YP_134542	128	2fvhA	60	2-99	1.66	-1.97
YP_326659	109	2fvhA	59	3-99	1.74	-1.64

* ModPipe Quality Score, ** a normalized DOPE (Discrete Optimized Protein Energy) score

```
1EF2_C        4 TPREKDKLLLLFTAALVAERRLARGLKLNYPESVALISAFIMEGARDG.[2].VASLMEEGRHVLTREQVMEGVPEMIPDI 80

Construct     4 TAKEQERLTVFTAAEVARRRKERGVPLNHPEAVAYISDWCIERGRDG.[2].VAEIRSGASKLLGREDVMDGVPEMIDMI 80

4UBP_A        5 NPAEKEKLQIFLASELLLRRKARGLKLNYPEAVAIITSFIMEGARDG.[2].VAMLMEEGKHVLTRDDVMEGVPEMIDDI 81

gi 17402589   4 EQREAEKLALHNAGFLAQKRLARGLRLNYTEAVALIAAQILEFVRDG.[3].VTDLMDLGKQLLGRRQVLPAVPHLLETV 81

gi 418162     4 TPREKDKLLLFTAGLVAERRLARGLKLNYPEAVALISCAIMEGARDG.[2].VAQLMSEGRTLLTAEQVMEGVPEMIKDI 80

gi 82702368   4 TPREKDKLQIFTAGLLAERRKARGLRLNYPEAVALITCAILEGARDG.[2].VAELMSEGRKVLTRADVMEGVPEMIPDI 80

gi 90591216   4 TPRESEKLLLHLAGELAAKRKARGLKLNYPETIAYISSHLLEAARDG.[2].VAELMNYGATLLTRDDVMEGIAEMIHDV 80

gi 6460755    4 TERERDKLLIFTAAQLARERRARGLKLNHPEAVALITAEVLEGIRDG.[2].VEDLMSFGAAILTPDDVLDGVPELIHEI 80

gi 2636191    4 TPVEQEKLLIFAAGELAKQRKARGVLLNYPEAAAYITCFIMEGARDG.[2].VAELMEAGRHVLTEKDVMEGVPEMLDSI 80

gi 14024886   4 TPREKDKLLIAMAAIVARKRLERGVKLNHPEAIALITDFVVEGARDG.[2].VAELMEAGAHVVTRAQVMQGIAEMIHDV 80
```

```
1EF2_C       81 QVEATFPDGSKLVTVHNPI 99
Construct    81 QVEPVFPDGTKLVTVHDPI 99
4UBP_A       82 QAEATFPDGTKLVTVHNPI 100
gi 17402589  82 QVEGTFMDGTKLITVHDPI 100
gi 418162    81 QVECTFPDGTKLVSIHDPI 99
gi 82702368  81 QVEATFPDGTKLVTVHNPI 99
gi 90591216  81 QIEATFPDGTKLVTVHSPI 99
gi 6460755   81 QVEGTFPDGTKLVTVHDPI 99
gi 2636191   81 QVEATFPDGVKLVTVHQPI 99
gi 14024886  81 QVEATFPDGTKLVTVHAPI 99
```

Figure 1. Multiple sequence alignment of urease constructs with functionally related domain sequences. (Shaded regions are showing alpha-gamma subunit interface)

Table 2. Data mining for searching metal-binding and active site similarity regions of urease models

NCBI Accession	Template ID	Metal-binding region	Active site*	Conservation Score
Urease alpha-subunit (PSSM-ID: 30031)				
YP_001190983	4ubpC	135-370	Ala150, Gly151, Phe152	4.375
Q18EB9	1a5lC	135-370	Gln364, Ala365, Met366	5.291
Q75ZQ5	1a5lC	135-370	Lys91, Arg92, Arg93	3.640
BAC84958	1a5lC	135-370	Gln362, Ala363, Met364	4.624
YP_134541	1a5lC	135-370	Gln362, Ala363, Met364	4.624
Q3IRZ5	1a5lC	135-370	Gln364, Ala365, Met366	5.264
Urease beta-subunit (PSSM-ID: 73201)				
CAJ53713	1ejxB	4-100	Gly98, Leu99, Val100	3.405
Q75ZQ6	4ubpB	4-100	Lys91, Arg92, Arg93	3.640
BAC84957	4ubpB	4-100	Lys91, Arg92, Arg93	3.640
YP_134540	4ubpB	4-100	Lys91, Arg92, Arg93	3.640
Q3IRZ6	1ejxB	8-104	Asp95, Arg96, Ile97	4.625
ABP95060	4ubpA	95-195	Asn97, Pro98, Ile99	0.843
YP_001190984	4ubpA	95-195	Asn97, Pro98, Ile99	0.843
Urease gamma-subunit (PSSM-ID: 63883)				
YP_659289	4ubpA	5-85	Arg26, Gly27, Val28	2.582
Q75ZQ4	2fvhA	5-85	Asp49, Gly50, Gln51	2.092
BAC84959	2fvhA	5-85	Asp49, Gly50, Gln51	2.092
YP_134542	2fvhA	5-85	Asp49, Gly50, Gln51	2.092
YP_326659	2fvhA	5-85	Gly72, Val73, Pro74	2.169

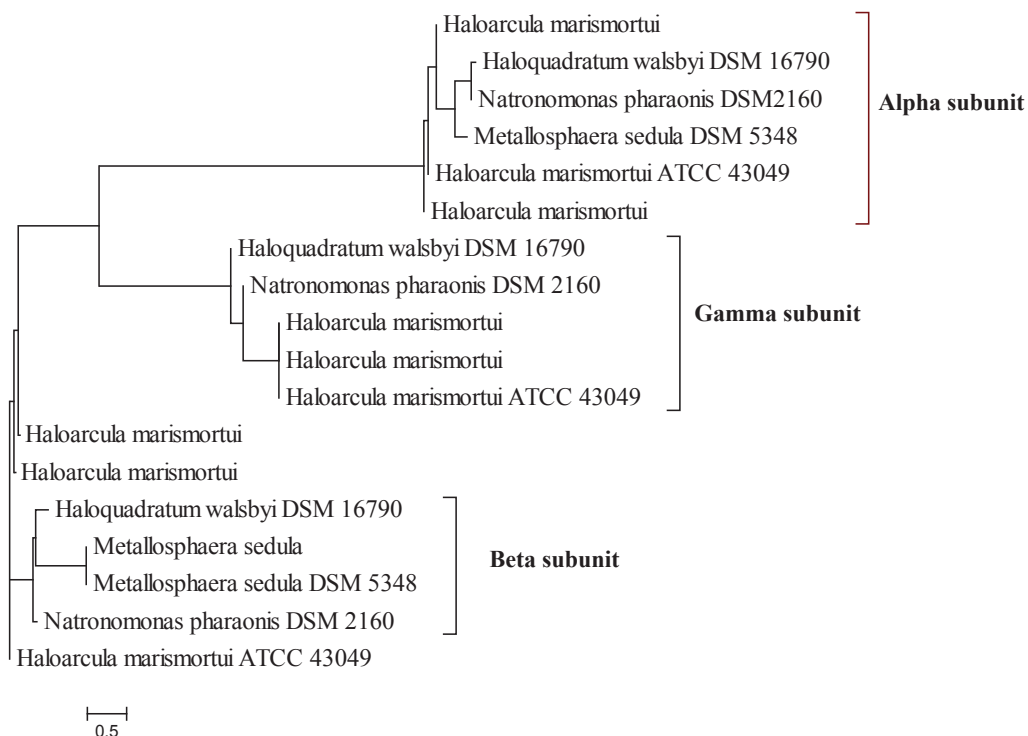*Active sites were predicted by ProFunc server

Figure 2. Phylogenetic tree constructed with NJ algorithm based on the urease sequences of archaea

and 0.1 RMSD. Using Ramachandran plot, 93.8%core, 4.9%allowed and 1.2%generally allowed regions (Figure 3). ERRAT2 program computed the overall structural quality (86.36%) of construct. Structural accuracy of construct was validated with Prove program by calculating Z-score RMS of that was 1.320 (Figure 4). Verify3d program resulted that 74.23%of the residues had an averaged 3D-1D score >0.2. The construct comprised of 52 H-bond donors and has 3 helix and 6 turns as secondary structural elements.

Molecular dynamic simulations of urease construct

A construct as selected from the above criteria was used to search lower energy conformers by molecular dynamic simulation by which 30 structural conformers were generated. As shown in Table 3, the top-five lowest

energy conforms were selected and had total energy around 2531-2135 kcal/mol. Torsion energy of each conformer has ranged from 197–215 kcal/mol. It also showed that electrostatic energies of all conformers was higher than vander Waals energies. As a result, energy conformer 1 of construct was chosen for further docking studies.

Molecular docking of substrate into urease construct

Urea, 2-propenamide (acrylamide), semicarbazide (aminourea), hydroxyurea and methylurea were chosen as substrates to construct in docking studies. We have shown that 6.78 mmol inhibition constant, -2.98 kcal/mol intermolecular energy and -2.54 kcal/mol internal energy when construct formed complex with semicarbazide (Table 4). Other substrates were
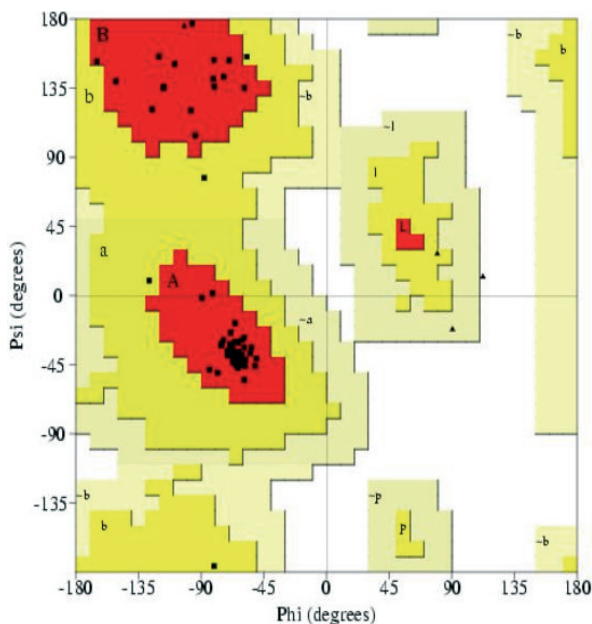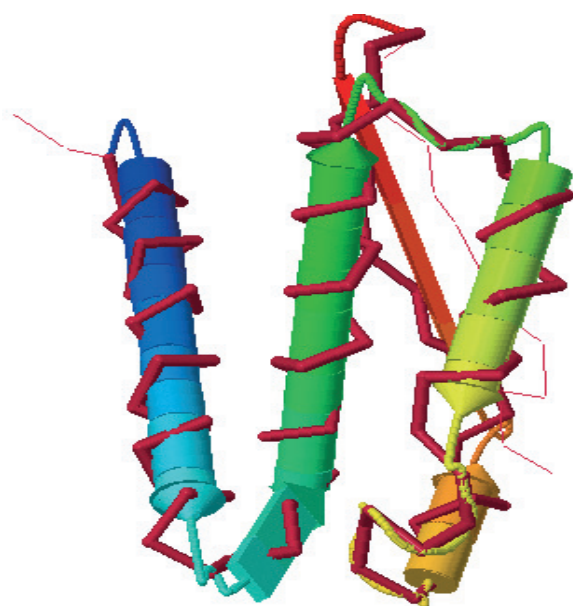
Construct
TAKEQERLTVFTAAEVARRRKERGVPLNHPEAVAYISDWCIERGRDGQSVAEIRSGASK
LLGREDVMDGVPEMIDMIQVEPVFPD
Template
TPHEQERLLLSYAAELARRRARGLRLNHPEAIAVIADHILEGARDGRTVAELMASGRE
VLGRDDVMEGVPEMLAEVQVEATFPD
Construct
LHHHHHHHHHHHHHHHHHHHHHHHLLLLLLHHHHHHHHHHHHHHHHHHHHHHLLLHHHHHH
HHHHLLLLHHHLLLLLHHHHLLLLLLLLLLLLLL
Template
LHHHHHHHHHHHHHHHHHHHHHHHHLLLLLLHHHHHHHHHHHHHHHHHHHHHHHLLLHHHHHH
HLLLLLLLHHHLLLLLHHHHLLEEEEEEEELL

Figure 3. Structure superimposition and quality analysis of urease construct. Figure in left represents the superimposition of construct vs 2FVH (A) and right represents Ramachandran plot for construct. Structural alignment of construct vs 2FVH (A) (H: Helix, L: Loop, E: Extended coil) is displayed in below figure.

Table 3. Molecular dynamic simulation data for top-five lower energy conformers of urease construct

Conformer	Total energy	Vander waals energy	Electrostatic energy	Torsion energy	Temperature K
1	-2532.08	-472.66	-2762.02	210.92	305.89
2	-2531.93	-488.67	-2725.28	197.40	311.50
3	-2531.91	-487.78	-2698.33	215.06	305.54
4	-2531.11	-482.35	-2745.80	204.13	302.35
5	-2531.05	-484.82	-2731.77	212.73	303.14

All of the molecular energies are expressed as kcal/mole.
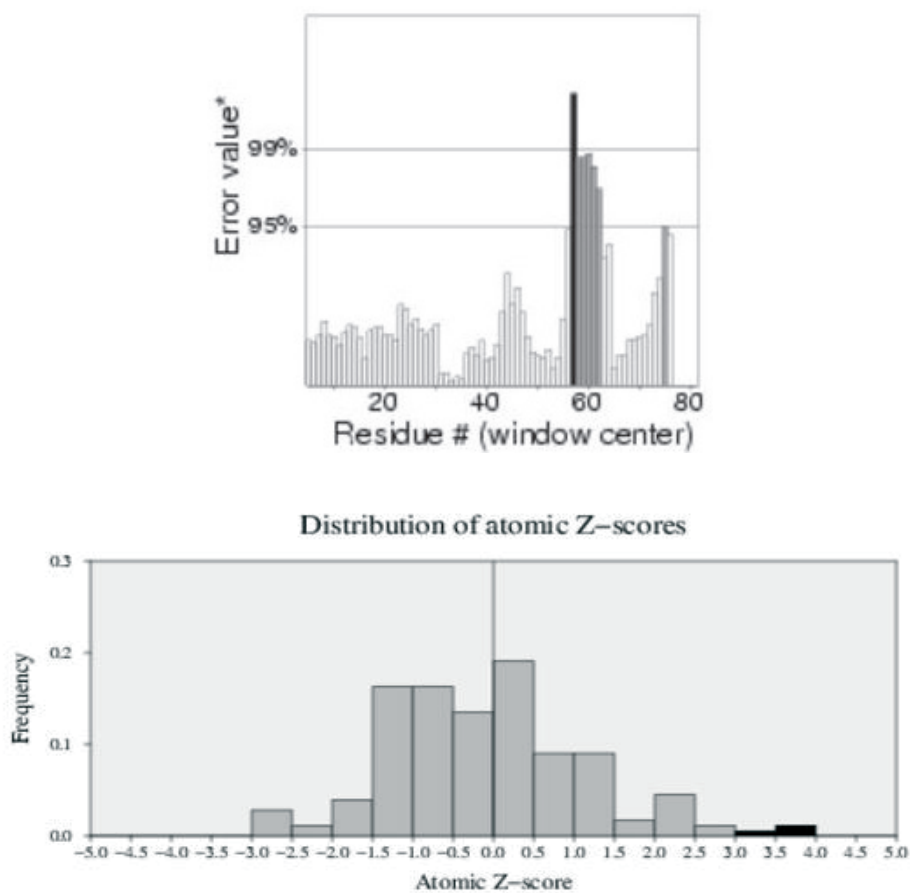
Figure 4. Structural validations of urease construct by ERRAT 2 Program (top) and by Prove Program (bottom)

Table 4. Substrate docking for predicting binding energy of substrate into urease construct

Substrate	RMSD (Å)	Binding energy (kcal/mol)	Inhibition constant (K_i mM)	Intermolecular energy (kcal/mol)	Internal energy (kcal/mol)
Urea	39.74	-2.69	10.74	-2.69	-2.41
2-Propenamide	41.42	-2.65	11.48	-2.92	-2.86
Semicarbazide	33.05	-2.96	6.78	-2.98	-2.54
Hydroxyurea	38.22	-2.47	15.41	-2.75	-2.54
Methylurea	33.76	-2.33	19.74	-2.60	-2.26

also formed complex with construct noticeably. As represented in Figure 5, atom N_2 of semicarbazide was H3-bonded to amino group of Gly50 of construct (2.97 Å) and atom bonded to carbonyl group of Gly50 (3.47 Å). Atom N_3 of semicarbazide was H4-bonded to amino group of Gln51 with interaction distance of 2.90 Å. Atoms N_1 and N_2 of semicarbazide was H-bonded to Glu55 residue with interaction distance of 2.79Å and 2.79Å, respectively.

Discussion

Most bacteria and archaea can grow under anaerobic conditions producing several enzymes that require nickel [26]. In this work, eight nickel-dependent enzymes are reported in NCBI to ensure their extensive role in biochemical processes of archaea. Urease is consisted of three subunits, alpha, beta and gamma, which can exist as separate proteins or can be fused on 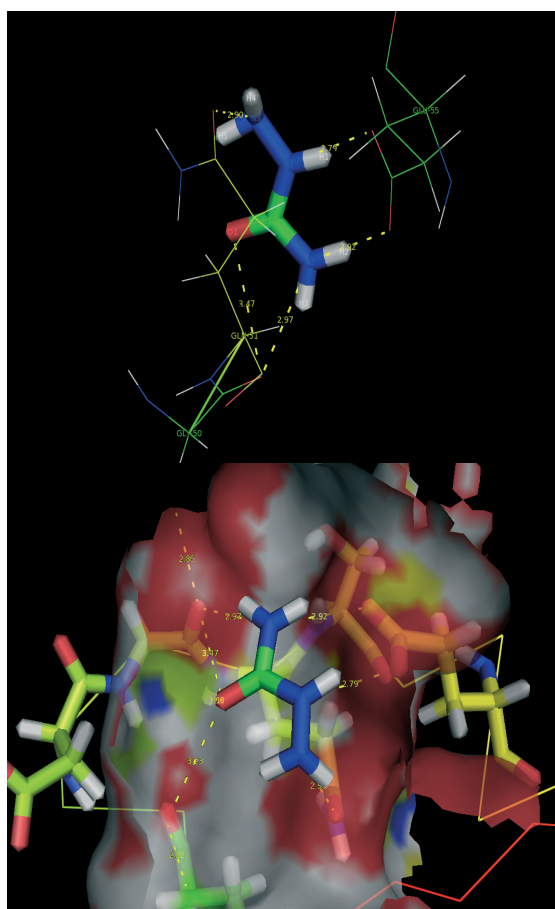a single protein chain [9,27]. A large alpha subunit is the catalytic domain containing an active site with a bi-nickel center complexed by a carbamylated lysine. The beta and gamma subunits are played a role in sub unit association to form the higher order trimers [25]. Since gamma subunit of urease has urease domain and active sites within nickel binding centre, we assumed that it could attribute in catalytic activity on respective substrates. Phylogenetic analysis of this study revealed that sequences of alpha and beta ureases were distantly related with gamma urease as they formed a separate clade apart from clade formed by archaeal gamma urease. *Metallophaera sudula* DSM 5348 showed more phylogenic relationship with urease from *H. marismortui* ATCC 43049. It implied that metal binding capacity and holotolerance of this enzyme are evolutionarily conserved within halobacteria and shared such features with organisms growing in metal containing environment.

As the results of docking studies, Gly50 and Gln51 was noted to form three H-bonds with semicarbazide, suggesting that this urease construct enable to catalyze it into form nitrile and water as similar to enzymes, nitrile hydratase, nitrilase, 3-cyanopyridine hydratase, NHase, L-NHase, H-NHase, acrylonitrile hydratase, aliphatic nitrile hydratase and nitrile hydro-lyase [28]. There was also a favorable interaction at Gly50 and Gln51 with urea-related substrates, and both amino acid residues have also reported as active sites of this enzyme. The second most binding energy (-2.65 kcal/mol) was computed when urease construct interacted with acrylamide. It has further supported us to ensure its catalytic efficiency like nitrile hydradase and acrylamide hydro-lyase on aminourea and acrylamide, respectively. Since, the proposed chemical reaction is that construct has catalytic competency to transform semicarbazide into nitrile, and acrylamide into acrylonitrile in aqueous environment (Figure 6).
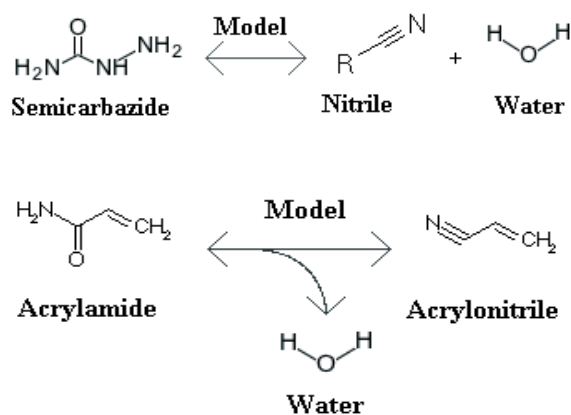
Figure 5. Molecular graphical representation of urease construct-semicarbazide complex. Top figure represents the molecular interaction view and bottom figure shows interaction view of substrate within binding pocket. The yellow dot lines denote the hydrogen bonds. All the amino acid residues which involved in molecular interaction are shown in line drawing and colored by residue types in which hydrogen is colored white, carbon green, oxygen red, nitrogen blue, and sulfur orange. Substrate is shown in stick in which carbon is colored tints, hydrogen gray, nitrogen blue, and sulfur orange. All interaction distances are represented as RMSD and expressed as Å.

Figure 6. Proposed chemical reactions catalyzed by designed enzyme. (Urease construct is proposed to convert semicarbazide (above) into nitrile and acrylamide into acrylonitrile (below) in aqueous environment)

A crystal structure obtained with urease from *Klebsiella aerogenes* indicated a trimeric (α3β3γ3) structure, but urease from *Helicobacter pylori* combined four of the regular six subunit enzymes in an overall tetrahedral assembly of 24 subunits (α12β12) [27, 29]. This supramolecular assembly is conferred additional stability for the enzyme, which functions to produce ammonia. Unlike structural complexity in natural urease, structure of urease construct obtained from this study has only 3 helices and 6 turns, it is assumed to confer its catalytic functions on urea-related compounds (aminourea and acrylamide). Nitrilases that hydrolyse organic nitriles to carboxylic acids and ammonia are a commercially interesting group of enzymes, as nitriles are important intermediates in the chemical synthesis of various products [30]. Nitrile-converting biocatalysts have considerable industrial interest as they can be used to treat toxic nitrile- and cyanide-containing wastes, and as agents for the synthesis of chemicals that have widespread applications [30-33]. Urease conductometric biosensors are used for a quantitative estimation of general water pollution with heavy-metal ions [34]. Therefore, this urease construct would probably take at least a few contributions for such applications.

Non-aqueous solvents, in which enzymes remain catalytically active, are emerging as versatile media for fundamental studies of enzyme structure and function. The enzyme activation is appeared to involve protection from the solvent in combination with transition-state stabilization, which may entail active-site hydration [35]. In some cases, thermodynamic equilibrium can be altered to favour dehydration over hydrolysis in organic solvents [36]. Kaul and Banerjee have studied the interaction of immobilized nitrile hydrolyzing biocatalyst with various water miscible organic solvents [37]. They found that the enzyme is the best functional in solvent concentration of 10–20%(v/v). Beyond the critical concentration, the conversion values dropped, probably due to enzyme inactivation upon contact with solvent. Nitrile biotransformation was the highly unstable nature of nitrilases, therefore, there were only a few successful biocatalytic processes utilizing nitrilases for large-scale production [36-38]. As the results of this study, urease construct is assumed to have similar catalytic function and transition-state stabilization with nitrilase.

Ilyina et al. have already demonstrated successfully the important role of artificially synthesized conservative amino acids in the principles of catalytic polypeptide properties by experimental evidence [3]. Molecular evolution-directed approach has already been reported for designing constructs of β-methylaspartate mutase from the sequences of *H. marismortui* [39], formyltetrahydrofolate ligase from *Haloquadratum walsbyi* DSM 16790 [40], sirohydrocholine cobalt chelatase and coenzyme F_{420} non-reducing hydrogenase from methanogens [41-42]. Since, the evolutionary conservation in sequence as well as structure would

make a major contribution in enzyme catalysis so that such conserved amino acid residues are accounted for designing enzyme. This designed enzyme has all promising features to perform chemical process in biotransformation reactions under artificial environment. Thus, the resulted urease construct provides an insight to appropriateness of using it in biotechnological processes and in green chemistry applications. For industrial applications, the gene sequence encoded for this construct should be synthesized and then expressed in either homologous host or heterologous host. The expressed protein should be analyzed for enzyme activity and substrate specificity in future.

Acknowledgements

References

[1] Monti D, Riva S. (2001) Natural and artificial microenzymes: Is it possible to have small and efficient biocatalyst? Biocat Biotrans 19:251–266.

[2] Luetz S, Giver L, Lalonde J. (2008) Engineered enzymes for chemical production. Biotechnol Bioeng 101:647–653.

[3] Ilyina AD, Ibarra-Coronado D, Gurumurthy K, Cerda-Ramarez F. (2006) Evidence of catalytic activity of polypeptides artificially synthesized from conservative amino acids. Moscow University Chemistry Bulletin 47:134–142.

[4] Lu Y. (2006) Metalloprotein and metallo-DNA/RNAzyme design: Current approaches, success measures and future challenges. Inorg Chem 45:9930–9940.

[5] Lu Y, Valentine JS. (1997) Engineering metal-binding sites in proteins. Curr Opin Struct Biol 7:495–500.

[6] Benson DE, Wisz MS, Hellinga HW. (2000) Rational design of nascent metalloenzymes. PNAS 97:6292–6297.

[7] Volkmer D, Krebs B. (1993) Dinuclear nickel (II) complexes as models for the active site of urease. J Inorg Biochem 51:511.

[8] Csiki C, Zimmer M. (1998) Computational design of biomimetic compounds: urease an example. J Mol Struct 442:11–17.

[9] Mobley HL, Hausinger RP. (1989) Microbial ureases: significance, regulation, and molecular characterization. Microbiol Rev 53:85–108.

[10] Lee MH, Mulrooney SB, Renner MJ, Markowicz Y, Hausinger RP. (1992) *Klebsiella aerogenes* urease gene cluster: sequence of *ure*D and demonstration that four accessory genes (*ure*D, *ure*E, *ure*F, and *ure*G) are involved in nickel metallocenter biosynthesis. J Bacteriol 174:4324–4330.

[11] Lee MH, Pankratz HS, Wang S, Scott RA, Finnegan MG, *et al*. (1993) Purification and characterization of *Klebsiella aerogenes* UreE protein: a nickel-binding protein that functions in urease metallocenter assembly. Prot Sci 2:1042–1052.

[12] Park IS, Hausinger RP. (1995) Requirement of carbon dioxide for in vitro assembly of the urease nickel metallocenter. Science 267:1156–1158.

[13] Moncrief MB, Hausinger RP. (1996). Purification and activation properties of UreD-UreF-urease apoprotein complexes. J Bacteriol 178:5417–5421.

[14] Hibbert EG, Dalby PA. (2005) Directed evolution strategies for improved enzymatic performance. Microb Cell Fact 4:29.

[15] Marshall SA, Mayo SL. (2001) Achieving stability and conformational specificity in designed proteins via binary patterning. J Mol Biol 305:619–631.

[16] Harbury PB, Plecs JJ, Tidor B, Alber T, Kim PS. (1998) High-resolution protein design with backbone freedom. Sci 282:1462–1467.

[17] Looger LL, Dwyer MA, Smith JJ, Hellinga HW. (2003) Computational design of receptor and sensor proteins with novel functions. Nature 423:185–190.

[18] Poole AM, Ranganathan R. (2006) Knowledge-based potentials in protein design. Cur Opin Struct Biol 16:508–513.

[19] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882.

[20] Tamura K, Dudley J, Nei M, Kumar S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599.

[21] Marchler-Bauer A, Bryant SH. (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res 32:327–331.

[22] Marchler-Bauer, A., Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. (2009) CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res 37:205–210.

[23] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

[24] Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, et al. (2003) Tools for comparative protein structure modeling and analysis. Nucleic Acids Res 31:3375–3380.

[25] Laskowski RA, Watson J D, Thornton JM. (2005) ProFunc: a server for predicting protein function from 3D structure. Nucleic Acids Res 33:89–93.

[26] Carrington PE, Al-Mjeni E, Zoroddu MA, Costa M, Maroney MJ. (2002) Use of XAS for the elucidation of metal structure and function: Applications to nickel biochemistry, molecular toxicology and carcinogenesis. Environ Health Pers 10:5.

[27] Jabri E, Carr MB, Hausinger RP, Karplus PA. (1995) The crystal structure of urease from *Klebsiella aerogenes*. Science 268:998–1004.

[28] Asano Y, Fujishiro K, Tani Y, Yamada H. (1982) Microbial degradation of nitrile compounds. Aliphatic nitrile hydratase from *Arthrobacter sp* J-1. Purification and characterization. Agric Biol Chem 46:1165–1174.

[29] Ha NC, Oh SC, Sung JY, Cha KA, Lee MH, et al. (2001) Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. Nature Struct Biol 8:505–509.

[30] Nigam VK, Khandelwal AK, Gothwal RK, Mohan MK, Choudhury B, et al. (2009) Nitrilase-catalysed conversion of acrylonitrile by free and immobilized cells of *Streptomyces* sp. J Biosci 34:21–26.

[31] Heald SC, Brandao PFB, Hardicre R, Bull AT. (2001) Physiology, biochemistry and taxonomy of deep-sea nitrile metabolizing *Rhodococcus* strains; Antonie van Leeuwenhoek 80:169–183.

[32] Trott S, Bauer R, Knackmuss HJ, Stolz A. (2001) Genetic and biochemical characterization of an enantioselective amidase from *Agrobacterium tumifaciens* strain d3. Microbiol 147:1815–1824.

[33] Brandao PFB, Bull AT. (2003) Nitrile hydrolyzing activities of deep-sea and terrestrial mycolate actinomycetes. Antonie van Leeuwenhoek 84:89–98.

[34] Burne RA, Chen YY. (2000) Bacterial ureases in infectious diseases. Microbes Infect 2:533–542.

[35] Clark DS. (2004) Characteristics of nearly dry enzymes in organic solvents: implications for biocatalysis in the absence of waterPhil. Trans R Soc Lond B 359:1299–1307.

[36] Layh N, Willetts A. (1998) Enzymatic nitrile hydrolysis in low water systems. Biotech Lett 20:329–331.

[37] Kaul P, Banerjee A, Mayilraj S, Banerjee U. (2004) Screening for enantioselective nitrilases: kinetic resolution of racemic mandelonitrile to (R)-mandelic acid by new bacterial isolates. Tet Assy 15:207–211.

[38] Kaul P, Banerjee CU. (2008) Predicting enzyme behavior in nonconventional media: correlating nitrilase function with solvent properties. J Ind Microbiol Biotechnol 35:713–720.

[39] Chellapandi P, Balachandramohan J. (2011) Molecular evolution-directed approach for designing β-methylaspartate mutase from the sequences of Haloarchaea. International J Chemical Mod 3:143-154.

[40] Chellapandi P, Balachandramohan J. (2011) Molecular evolution-directed approach for designing archaeal formyltetrahydrofolate ligase. Turk J Biochem 36:122-136.

[41] Chellapandi P, Balachandramohan J. (2012) Cobalt and nickel-containing enzyme constructs from the sequences of methanogens. Biopolymers Cell 28 :68-74.

[42] Chellapandi P. (2011) Current scenario of computer-aided metalloenzymes designing. Biopolymers Cell 27:432-435.