Cobalt- and Nickel-containing enzyme constructs from the sequences of methanogens

P. Chellapandi, J. Balachandramohan

Department of Bioinformatics, School of Life Sciences, Bharathidasan University
Tiruchirappalli-620024, Tamil Nadu, India
pchellapandi@gmail.com

Introduction. Enzymatic or microbial transformations are environmentally friendly, clean processes and useful means for obtaining biologically important compounds. Enzymes from nature rarely have the combined properties necessary for the industrial fine chemical production that will be present over the course of a manufacturing process. Enzymes from extremophilic organisms may provide useful biocatalysts and may be even more valuable for biotransformation reactions. A range of methanogenic archaeon enzymes application and usage of the organisms themselves in biotechnology are more restricted. At present, many studies explore the identification of novel enzymes from methanogens for the industrial application [1, 2].

Metals are tightly bound in the active sites of metalloenzymes to bring the chemical activity. The metalloenzymes are potentially very good models for designing chemical catalysts. The first step towards designing a working catalyst is, therefore, knowledge on the structure of the enzyme and its active site [2, 3]. Evolutionary conservation in sequence and structure would make contribution in enzyme catalysis. Such conserved amino acid residues are major concern in designing metalloenzymes.

Upon designing enzyme constructs, the molecular evolution of enzymes is playing a crucial role in biological systems when enzymes are in action [4, 5]. Thus, the present study describes the application of a molecular evolutionary hypothesis to design metalloenzymes constructs from conserved domains in the sequences of methanogens encompassing metal- and substrate-binding sites.

Materials and methods. Evolutionary conservation analysis of metalloenzymes. Complete protein sequences of archaeal cobalt and nickel-containing enzymes were retrieved from GenPept of NCBI. Conser-
ved domains architecture of these sequences was searched in NCBI-Conserved Domain Database (CDD). The metal binding sites of the sequences were identified from availed PDB structures using the corresponding PSSM ID in NCBI-CDD. A position-specific score matrix prepared from the underlying conserved domain alignment was compared with the query sequences. Multiple sequence alignment of the selected sequences was performed by ClustalX 2.0 software [6], in which aligned sequences were manually inspected to delete the low scoring sequences. After that, the neighbor joining phylogenetic tree was constructed by MEGA 4.0 software [7] with 1000 bootstraps values.

**Molecular modelling and enzyme designing.** Homology modelling was carried out by ModWeb, an automatic comparative protein modelling server, from query sequences obtained from methanogens [8]. A suitable PDB template for homology modelling was identified with PSI-BLAST tool [9] by searching against PDB database. ProFunc server [10] was used to predict the corresponding function sites of models. The catalytic domains of every protein model were compared with the crystallographic protein structures. Amino acid residues encompassing the sites for metal-binding and substrate-binding regions were removed from atomic coordinates of a protein model and subjected to further homology modelling. Prime program in Maestro software package (Schrodinger Inc.) was used to build the homology models from the above selected regions and then evaluated with Structural Analysis and Verification Server (SAVS) (http://nihserver.mbi.ucla.edu/SAVES/). The best scoring models were superimposed on the corresponding PDB template by DALITE server (http://ekhidna.biocenter.helsinki.fi/dali_lite/start). Standard dynamics simulation cascade module in Discovery Studio software with CHARMM force field, steepest descent and adopted basis Newton-Raphson algorithms was used for generating structural conformers of each model. Conformers were typically created for molecules that have a small number of alternate locations for a small subset of the atoms. Distance constraint of each model was fixed between N-terminal to C-terminal, and dihedral restraint was started from C to Ca (Φ) of first amino acid residue and Ca to N (Ψ) of second amino acid residue until the last amino acid residue in a molecular dynamic ensemble.

**Molecular docking of enzyme construct-substrate complex.** Substrate structures were retrieved from KEGG database with SIMCOM tool (http://www.genome.jp/tools/simcomp) and then molecular formats prepared as PDB format for molecular docking studies. Homology models resulted from Prime program was prepared with charges for molecular docking studies. Both substrate and protein were treated with AMBER force field implemented in AutoDock 4.0. software. Ligand Fit docking was conducted by AutoDock software with Genetic Algorithm and computed inhibition constant, binding energy and other molecular forces of resulted docking models. Active site residues in enzyme construct as in energy grid were allowed to interact with substrate. After docking, enzyme-substrate complex was solvated with explicit periodic boundary solvation model. The quality of each docking model was checked by AutoDock scoring system and graphically represented by PyMOL software.

**Results and discussion.** Designing sirohydrocholine cobalt chelatase. Cobaltochelatase, methionine synthase, methionine aminopeptidase, methylmalonyl-CoA mutase and methyl aspartate mutase are cobalt-dependent enzymes identified in the genomes of archaea. Among these, sirohydrocholine cobaltochelatase (SHCCC) and anaerobic chelatase (ACC) are present in methanogens. The limited structural information is available for cobalt chelatase in the PDB to date. The sequence identity of SHCCC is ranged from 30–46 % with PDB templates. Enzyme sequences under accession numbers YP_001098022 (construct 1) from *Methanococcus maripaludis* C5 and YP_001030522 (construct 2) from *Methanocorpusculum labreanum* Z have significant modelling scores, and encompassed the shortest metal binding sites incorporating functional region (Supplementary). Aligned SHCCC sequences have also more conservation at the substrate- and metal-binding regions (Fig. 1). Construct 1 is evolutionarily related only to closely related species of the same genera and its function may conserve in the particular sequence position and/or domain. SHCCC in *M. labreanum* Z shows more evolutionary relationship with SHCCC of *M. maripaludis* C5 (bootstrap value 987–1000) than with other archaea. The different species of *Methanosarcina* genus are clustered with the members in the genus of *Methanococcus* in the first clade, and then both of them
formed second clade having organisms belonged to the groups of halophilic and sulfur reducing archaea (Fig. 2). As the results of phylogenetic analysis, evolution-based approach provides a high likelihood to design metalloenzymes.

Construct 2 is slightly better than construct 1 as it has the lowest binding energy (−9.79 kcal/mol). Both enzyme constructs are opted for catalytic conversion uroporphyrin I and precorin-2-dihydro sirohydrochlorin to the respective products (Table). The most important interactions contributing to the high binding affinity are six H-bonds between the carbonyl group of sirohydrochlorin and the side chains of His09, Leu70, Arg12 and Ser16 in construct 1. The possible interaction sites are: O2-His09 (3.10 Å), O15-Leu70 (2.93 Å), O7-Ser16 (2.96 Å), O11-Arg12 (3.19 Å), O5-Arg12 (2.78 Å), and O5-Arg12 (3.38 Å). Among these, Arg12 forms three H-bonds to confer higher binding affinity. Six H-bonds are formed between sirohydrochlorin and construct 2 at atomic positions of H3-Ser13 (2.52 Å), O6-Ser13 (3.08 Å), O1-Arg14 (3.00 Å), O5-Arg14 (2.98 Å), O12-Lys19 (2.85 Å), and O16-Lys19 (2.76 Å) (Supplementary).

Designing coenzyme F_{420} non-reducing hydrogenase. Archaeal genomes have consisted of 6 cobalt containing enzymes include coenzyme F_{420} hydrogenase, methyl coenzyme M reductase, hydrogenase maturation protease, carbon monoxide dehydrogenase, rubredoxin, urease, acetyl-CoA decarbonylase/synthase complex. Coenzyme F_{420} hydrogenase has two types of enzyme modules, such as coenzyme F_{420}-reducing hydrogenase and coenzyme F_{420}-non-reducing hydrogenase (CFNRH). As similar to SHCCC, no crystallographic structures of these enzymes in the PDB are opted for enzyme designing. The protein sequence of Methanosarcina acetivorans C2A under accession number NP_616085 (construct) is suited on designing CFNRH construct (Supplementary). As compared to coenzyme F_{420}-reducing hydrogenase and non-reducing hydrogenase, hydrogenase maturation protease is distantly related for its function. CFNRH of this study is phylogenetically resembled to coenzyme F_{420}-reducing hydrogenase of other methanogens (Fig. 3). It pointed out that coenzyme F_{420}-reducing hydrogenase (hysF) and hydrogenase maturation protease (hycI) are distinctly produced monophyletic clades in the phylogenetic tree. There are separate sub-clusters for the species of all genera including Methanococcus, Methanosarcina, Methanospirillum and thermophilic archaea in
which many of the hydrogenases belonging to *Methanococcus* genus. It may functionally diverge into NiFe-hydrogenase maturation protease of *Candidatus Korarchaeum cryptofilum* OPF8 (Fig. 4).

Enzyme construct has shown strong interactions with reduced FMN. It is also supported by computing inhibition constant (38.58 µM) of reduced FMN as shown in Table 1. Intermolecular energy of construct-reduced coenzyme F$_{420}$ complex is –12.26 kcal/mol. A strong binding affinity is observed in enzyme construct-reduced FMN wherein two H-bonds formed between the side chains of Lys80 and Ala86 and reduced FMN. The possible interaction sites are: O1-Lys80 (2.97 Å) and O6-Ala86 (2.88 Å) (Supplementary).

The single conserved domain (CbiX_SIR_B_N of chelatase class II superfamily) is detected in SHCCC sequences of *M. maripaludis* and *M. labreanum* Z that may catalyze biochemical reactions. A divergent evolution occurred among cobalt and nickel metabolisms due to more ancestral similarities of cobaltchelatase as reported earlier [11]. Cobaltochelatase of archaeal domain has shown more phylogenetic similarity with CbiX of *Bacillus megaterium* [12]. Similarly, the phylogenetic analysis in this study revealed a phylogenetic resemblance between methanogens and *B. megaterium* for SHCCC sequences. It also closely related to halophilic archaea. CbiX tolerates large sequence rearrangements in four out of the nine loop regions of the SHCCC.
These rearrangements may serve as starting points for the evolution of new functions (binding of new substrates) [13]. It suggested that conservations of enzyme function are not deviated at sequence level, and its catalytic function evolved from the structural features, particularly loop regions of these constructs. Thus, the present approach is more convenient for designing an enzyme even from its sequence in appropriate manner.

Glu264 and His183 are mechanistically the most critical residues for catalytic activity and form the active site in conjunction with residues that selectively bind the tetrapyrrole substrate in accordance with the earlier reports [14, 15]. The results from molecular mechanics studies implied that conformational free energies of our enzyme constructs are more stable. Hence, the structures of these constructs have to show low degree of freedoms to recognize sirohydrochlorin specificity, which is rather than other tetrapyrole substrates (Supplementary).

Removal of two protons followed by insertion of the cobalt ion induces planarity into macrocycle, resulting in its exclusion from the active site [16]. Thus, sirohydrochlorin is bent such that two pyrrole nitrogens are orientated towards the active site histidine for its binding with SHCCC. Similarly, sirohydrochlorin stably interacts with the side chains of His09, Leu70, Arg12 and Ser16 in construct 1 and Ser13 and Arg14 in construct 2. A strong interaction is found between construct 1 and sirohydrochlorin in the side chain of Arg12. At Ser13 and Arg14 residues, construct 2 forms a complex with sirohydrochlorin. It indicates that either Ser13 or Arg12 could serve as active region providing catalytic function to these constructs. Towards the C-terminal end of CbiX, the histidine rich region is likely to act as a cobalt store or could interact directly with the cobalt transport system to allow direct transfer of the metal ion from the import system to its site of action. Construct 1 has more numbers of histidine residues, some of which contributed for H-bonding with sirohydrochlorin in accordance to the earlier report [17].

There is no common agreement on the catalytic mechanism of Ni-Fe hydrogenase. The sequences of CFNRH from M. acetivorans consisted of H2MP domain assigning function to hydrogenase specific C-terminal endopeptidases. Phylogenetic analysis in this study also supported its functional divergence with NiFe-hydrogenase maturation protease of Candidatus Korarchaeum cryptofilum OPF8. H2MP domain may be originated from subsequent functional divergence during evolution. The maturation of FeNi-hydrogenases includes the formation of nickel metallocenter, proteolytic processing and assembly with other subunits. The hydrogenase maturation endopeptidases are responsible for the proteolytic processing, liberating a short C-terminal peptide by cleaving after His or Arg residue. This cleavage is nickel dependent. Similarly, in this study a strong binding affinity is observed between individual atoms of reduced FMN and the side chains of enzyme construct. It reflects the enzyme mimics as similar to hydrogenase maturation endopeptidases by cleaving His87 or Arg84 residue in the C-terminal. Based on DFT studies, Pavlov et al. proposed a mechanism stating the resting state of the dinuclear cluster is Ni (II) Fe (III); H2 first binds to Fe in the form of a molecular hydrogen complex, which then undergoes heterolytic splitting [18]. Hydride transfer
to Fe and proton transfer to the adjacent Cys thiolate ligand is accompanied by decoordination of the protonated Cys thiol from Ni while remaining bound to Fe. The cyanide ligand on Fe simultaneously binds with the Ni atom in a bridging binding mode. After the H$_2$ dissociation, the hydride bound to Fe can then be transferred to Ni which should be a necessary preliminary for subsequent H$^+$ or electron transport. Accordingly, oxidation of molecular hydrogen could be mediated by CFNRH construct on reduced FMN. A strong phylogenetic correspondence of these enzyme constructs only to other closely related methanogens implied its evolutionary conservation.

Similar approach has already been reported for designing β-methylaspartate mutase [3] and formyltetrahydrofolate ligase [5] from the sequences of archaea [19]. Similar to that, we have not done any alternation in native amino acids position or replacing residues, and, thus, it is directly based on the conserved domain of metalloenzymes.

**Acknowledgements.** The first author is thankful to the University Grants Commission, New Delhi, India, for financial assistance (32-559/2006) to carry out the work.

**II. Челпанаді, Дж. Бахананіранганан**

Кобальт- і нікель-вмісні ферментні конструкції із послідовностями метаногенів

**Резюме**

**Мета.** Використати консерваційні домени, отримані із послідовностей металлопротеїнів, для конструювання ферментів. Особливу увагу приділяємо металлоферментам, оскільки вони мають важливу еволюційну значущість. **Методи.** Теорія молекулярної еволюції, методи молекулярного моделювання і молекулярного докінців. **Результати.** Конструювання кобальт-вміснії кобальтової хелати і нікель-вмісного конферменту F$_{\text{m}}$, нередукційної гідрогенази із консерваційних доменів, ототожнюючи металічні i субстратик углеводорідній нікелевий сайт, стосовно до основної теорії молекулярної еволюції. Зроблено припущення, що існує будь-який фермент містить у своєму каталітичному сайті стійкий або ідентичний консерваційний домен, то конструкцій може бути прийнята подібна каталітична активність. Використовуючи цей підхід для створення ферментної конструкції, потребно зможе
Chellapandi P., Balachandramohan J.


Received 10.05.11