

Full structure building and docking of NifS from extremophile *Acidithiobacillus ferrooxidans*

LIU Yuan-dong(刘元东), QIU Guan-zhou(邱冠周), WANG Hai-dong(王海东),
JIANG Ying(蒋莹), ZHANG Cheng-gui(张成桂), XIA Le-xia(夏乐先)

School of Minerals Processing and Bioengineering, Central South University, Changsha 410083, China

Received 29 August 2007; accepted 25 November 2007

Abstract: The gene *iscS-2* from extremophile *Acidithiobacillus ferrooxidans* may play a crucial role in nitrogenase maturation. To investigate the protein encoded by this gene, a reliable integral three-dimensional molecular structure was built. The obtained structure was further used to search binding sites, carry out the flexible docking with cofactor pyridoxal 5'-phosphate(PLP) and substrate cysteine, and identify its key residues. The docking results of PLP reveal that the residues of Lys203, His100, Thr73, Ser200, His202, Asp177 and Gln180 have large interaction energies and/or hydrogen bonds fixation with PLP. The docking results of cysteine show that the amino group in cysteine is very near His100, Lys203 and PLP, and the interaction energies for cysteine with them are very big. These identified residues are in line with the experimental facts of NifS from other sources. Moreover, the four residues of Asn152, Val179, Ala102 and Met148 in the PLP docking and the two residues of Lys208 and Ala102 in the cysteine docking also have large interaction energies, which are fitly conserved in NifS from all kinds of sources but have not been identified before. According to these results, this gene encodes NifS protein, and the substrate cysteine can be effectively recruited into the active site. Furthermore, all of the above detected key residues are directly responsible for the binding and/or catalysis of PLP and cysteine.

Key words: bioleaching; NifS; *Acidithiobacillus ferrooxidans*; homology modeling; molecular dynamics; docking; pyridoxal 5'-phosphate(PLP); cysteine

1 Introduction

Acidithiobacillus ferrooxidans, a gram-negative and chemolithotrophic bacterium, is one of the most broadly used bacteria in industrial mineral processing to extract metals such as copper, zinc, uranium, lead, gold and nickel from their insoluble sulfide minerals[1]. The bioleaching environments for its survival are extreme acidity and toxicity[2]. One may wonder how life is possible in such extreme environments. In fact, *A. ferrooxidans* can not only survive in these poisonous surroundings[3–4], but also take advantage of these conditions to gain energy for growth through oxidizing ferrous ion and reduced sulfur compounds as well as to convert the insoluble metal sulfides to their soluble metal sulfates[5]. This capability is therefore utilized to extract metals from minerals by industry. At the same time, *A. ferrooxidans* also fixes atmospheric N₂ to provide cellular N[1]. The nitrogen availability of this bacterium

may limit its growth and then affect the efficiency of bioleaching operations. So the study of nitrogen fixation system of *A. ferrooxidans* is of both fundamental and applied interest.

Nitrogen-fixing bacteria possess a nitrogenase that catalyzes the reduction of atmospheric nitrogen to ammonia (nitrogen fixation). The nitrogenase is a complex [Fe-S] enzyme containing two component, Fe protein and MoFe protein[6–7]. The studies on nitrogenase maturation revealed the insights concerning a pathway for [Fe-S] cluster biogenesis[8]. In this processing, a consortium of proteins is required for the synthesis and insertion of the associated metalloclusters. Among these consortium proteins, the NifS functions mainly as sulfur donor for iron-sulfur cluster formation, in which NifS gets sulfide from cysteine, and then works together with NifU to produce [Fe-S] cluster precursors destined for nitrogenase maturation[9]. NifS is a cysteine desulfurase, a pyridoxal 5'-phosphate(PLP)-dependent enzyme that catalyzes the conversion of L-cysteine to L-

alanine and sulfan sulfur via the formation of a protein-bound cysteine persulfide intermediate on a conserved cysteine residue[10–11]. Several molecular structures of cysteine desulfurase from various sources have been resolved by X-ray crystallography[12–15]. However, in all of those structures, a residue segment containing the active site cysteine and some residues in N-terminal and C-terminal regions cannot be directly resolved because these regions are disordered and have poor electron density. Moreover, none of them contains substrate cysteine.

It was reported that the genome sequence of *A. ferrooxidans* ATCC 23270 includes a gene *iscS-2* concerning nitrogenase maturation, which may encode the NifS protein. But till now, there are no experimental and theoretical efforts made in the protein encoded by this gene and its three-dimensional(3D) structure remains to be elucidated. Furthermore, it is an obstacle for directly obtaining the structure information of substrate-enzyme complex because substrate and enzyme react to provide product in a split second whereas the structure resolutions such as X-ray crystallography generally take many hours. To the best of our knowledge, homology modeling and molecular docking *in silico* are efficient methods for the 3D structure construction of protein and the interaction investigation of substrate-enzyme, respectively[16]. The detailed 3D structure, substrate-enzyme interaction information and identified key residues of this protein are, however, essential for guiding site-directed mutagenesis experiments and understanding the catalytic mechanism and function in the unique physiology of *A. ferrooxidans*.

In this work, with homology modeling techniques and molecular dynamics simulations, an integral 3D molecular model of NifS from *A. ferrooxidans* (afNifS, encoded by the *iscS-2* gene) was created, refined and assessed. The obtained structure was further used to search the bind sites, carry out the flexible docking with cofactor PLP and substrate cysteine, and identify its key residues.

2 Computation and methods

The primary amino acid sequence of the afNifS was retrieved from the *A. ferrooxidans* ATCC 23270 genome from the Institute for Genomic Research(TIGR). All simulations were performed on the Dell Precision 470 workstation with Redhat Linux system using Insight II software package developed by Accelrys Software Inc.

2.1 Three dimensional molecular model building

The Homology module[17] was used to build the initial model of afNifS. Firstly, a sequence similarity

search on Protein Data Bank(PDB) by BLAST program was carried out to find related protein structures as templates. Then, Modeler program was performed to build the 3D structure of afNifS. In order to construct the residue segment containing the active residue cysteine, a loop search algorithm over the databank of known crystal structure was used. The residues at the N-terminus and C-terminus were generated through End-Repair program.

The initial model was further improved by using Discover_3 module. During this procedure, an explicit solvent model with a 2 nm water cap from the center of mass of afNifS was used and the consistent-valence forcefield(CVFF) was selected. After performing 800 steps of conjugate gradient(CG) minimization, molecular dynamics(MD) simulation was carried out to check their stability via performing 70 ps simulations at a constant temperature of 298 K. Finally, a conjugate gradient energy minimization(EM) of the full solution model was performed until the root mean square(RMS) gradient energy was lower than 42 GJ/(mol·m).

The final structure was assessed by using Profile-3D [18] and ProStat[19] programs.

2.2 Binding sites analysis

In order to investigate the interactions among protein afNifS, cofactor PLP and substrate cysteine, the binding pockets are defined as subsets containing residues, in which any atom is within 0.3 nm from PLP and cysteine, respectively. The binding pockets of the modeled structure were identified through locating cavities by using Binding-Sites module. Then, combined with the search results, the conserved residues in the family of NifS and the experimental results of other analogs to afNifS, and the binding sites of afNifS for PLP and cysteine were predicted. Those results were used to guide the following docking experiments.

2.3 Docking cofactor and substrate into binding pockets

The 3D structures of cofactor PLP and substrate cysteine were built by Builder program and their geometries were further optimized through Discover_3 program. For tracking the interaction modes among enzyme, cofactor and substrate, the molecular docking among them was performed by using docking program Affinity.

Firstly, the cofactor PLP was hosted into the binding pocket of afNifS apoenzyme under the guide of the above identified site. During the procedure, the best binding conformation of the ligand to the receptor was automatically searched by Affinity program based on the

energy of the total complex. The potential function of the complex was assigned by using the CVFF. The cell multipole approach was used for nonbonding interactions. To account for the solvent effect, the centered ligand-receptor complex was solvated in a sphere of water molecules with a radius of 3 nm. The final docked complex was selected according to the criteria of interaction energy combined with geometrical matching quality.

Secondly, the substrate cysteine was docked into the active site of afNifS holoenzyme under the guide of the above identified site. Since PLP is covalently bonded with enzyme NifS in nature, the PLP bonded to enzyme was manually done and then the MD and EM processing was performed to remove the tensional structures before this docking. The docking procedure of cysteine was the same as that of PLP.

3 Results and discussion

3.1 Molecular structure modeling of NifS from *A. ferrooxidans*

The results of search on PDB for the protein sequence of afNifS showed that the crystal structure of a cysteine desulfurase from *E. coli* (PDB code 1P3W)[12] had the best sequence identity (41.56%) with it, so 1P3W was used to model the 3D structure of afNifS. However, the structures of a segment of 12 residues containing the active site cysteine and some residues in C-terminal region in 1P3W are disordered and cannot be directly solved due to the poor electron density in these regions. The situations for all the solved structures analogous to NifS are the same as 1P3W. This makes an obstacle to further insight into the function of these enzymes. Fortunately, the homology modeling techniques in Insight II can be used to settle such problems.

The sequence alignment between the reference

protein and template protein is shown in Fig. 1. Based on the alignment, through the procedures of homology modeling, the initial molecular model of afNifS was completed.

The initial structure was further refined. During the procedure of molecular dynamics simulation, the potential energy of the full solution model fell rapidly in the first 40 ps, and then decreased with very low deviation between two steps, and the dynamics simulations tend to equilibrium at 50 ps (Fig.2). So the conformation at 70 ps was chosen to further minimize as the final 3D structure.

When the final structure is accessed by ProStat, 86.2% of the Φ - Ψ angles in the model lie in the core region of the Ramachandran plot, and compared to 1P3W, the percent of backbone Φ - Ψ angles is 85.6%. When the cutoff used for significant difference is six standard deviations from the reference values, no significant differences of structure features in the modeled protein are observed. When being checked by Profile-3D, the self-compatibility score for this structure is 181.95, which is higher than the low score (83.54) and close to the top score (185.65). The compatibility scores for each residue in the modeled structure are all above zero (Fig.3), which corresponds to acceptable side chain environments. So, both of the above results from Profile-3D and ProStat programs indicate that the modeled structure is reliable.

The modeled afNifS monomer consists of 402 amino acids, and has a relative molecular mass of 44 170 by calculation. It can be subdivided into a small and a large domain (Fig.4). The small domain is composed of the N-terminal residues 1–14 and the C-terminal residues 260–402, which has six helices, two sheets and seven turns. The larger domain contains residues 15–259, which has eight helices, seven sheets and fourteen turns. The overall structure feature of this protein is typical for

```

1P3W ( 2) KLP IYLDYSATTPVDPRVAEKMMQFMTMDGTFGNPASRSHRFGWQAEAAVDIARNQIADL (61 )
afNifS( 1) mssaymdnnattrvdpqvlletilpyfs--eyygnpssm-hsfggdvgrkrleearasvqal (57 )
      |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1P3W (62) V G A D P - R E I V F T S G A T E S D N L A I K G A A N F - Y Q K K G K H I I T S K T E H K A V L D T C R Q L E R - E G (118)
afNifS(58) l g v e h d s e i i f t s c g t e s d n t a i l s a l a i q p s r r - - e i i t s v v e h p a i l s l c d h l e r h e g (115)
      |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1P3W (119) F E V T Y L A P - Q R N G I I D L K E L E A A M R D D T I L V S I M H V N N E I G V V Q D I A A I G E M C R A R G I I Y (177)
afNifS(116) y k v h y i - p v d r l g r i s m e a y a a l s e q v a v v s m m y a n n e t g t i f p v e s m a r m a k e k g a l f (174)
      |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1P3W (178) H V D A T Q S V G K L P I D L S Q L K V D L M S F S G H K I Y G P K G I G A L Y V R R K P R V R I E A Q M H G G G H E R (237)
afNifS(175) h t d a v q a v g k i a l n m q e s a i d n l s l s g h k l h a p k g v g a l y v r r g t r y r p - - m l r g g h q e r (232)
      |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1P3W (238) G M R S G T L P V H Q I V G M G E A Y R I A K E E M A T E M E R L R G L R N R L W N G I K D I E E V Y - - - L N G D L E (294)
afNifS(233) g r r a g t e n t t g i a l g l a a d m a l e h m e y e n t v v r a l r d r l e a g i - - l a q v p y a f v n g d p e (290)
      |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1P3W (295) H G A P N I L N V S F N Y V E G E S L I M A L K - - D L A V S S G S A - - - - - E P S Y V L R A L G L N D E L A H S S (352)
afNifS(291) h r l p n t s s i s f e y i e g e a i l m l l s r v g i a a s s g s a c t s g s l e p s h v l r a m d i p y t a a h g t (350)
      |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1P3W (353) I R F S L G R F T T E E E I D Y T I E L V R K S I G R L R D L S P L W E M Y K Q (392)
afNifS(351) i r f s l s r e t t e a e v d w v i a q t p p i v e q l r r l s p y w g a n g p i l t d n a f a p v y a (402)
      |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

```

Fig.1 Sequence alignment of afNifS with 1P3W

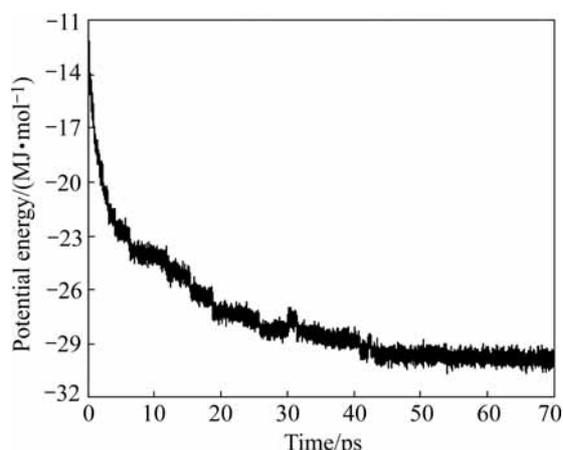


Fig.2 Variation of total energy during 70 ps of MD on afNifS (Total energy is averaged over 0.1 ps interval)

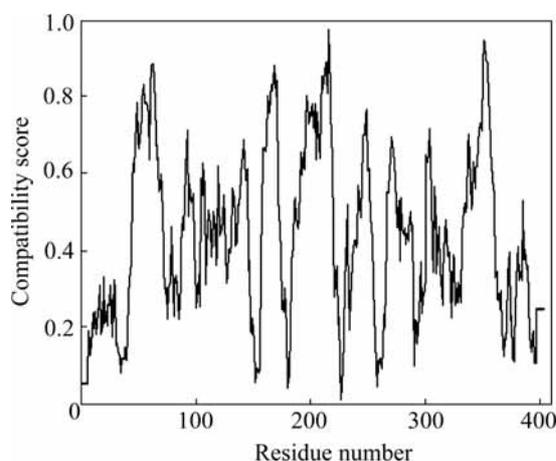


Fig.3 3D profiles of verified results of afNifS model (Residues with positive compatibility score are reasonably folded)

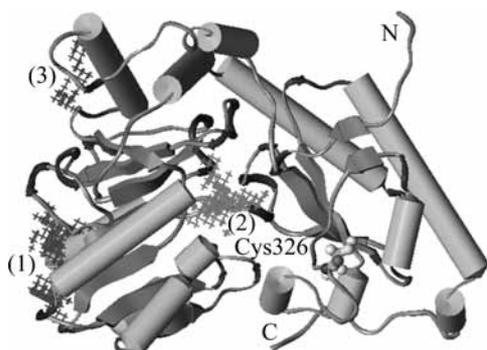


Fig.4 Final 3D-structure and possible binding sites of afNifS (N-terminal and C-terminal of this protein are noted N and C, respectively; active residue cysteine is represented by balls and sticks; possible binding sites are represented by filling their space with intersections and noted 1, 2 and 3, respectively)

NifS protein. Moreover, the structures of N-terminal and C-terminal of this protein are random coils and disorders. The structure of the residue segment containing the active site residue Cys326 randomly bulges out in the

orientation of the large domain. These situations are also in line with the case that cysteine desulfurase is disordered in these regions.

3.2 Identification of PLP binding and cysteine active regions

NifS is a PLP-dependent cysteine desulfurase. In the course of the catalytic action, a substrate, ketimine intermediate is produced from L-cysteine and PLP, and the sulfur atom of this intermediate is attacked by the catalytic cysteine residue, resulting in the formation of a cysteine persulfide residue at the active site of the enzyme[12–15]. The substrate derived sulfur atom of the cysteine persulfide residue is then transferred to various sulfur-containing biomolecules. It is reported that the cysteine desulfurase has a PLP-binding pocket and this pocket is involved in cofactor binding and substrate catalysing[12–15]. In this pocket, there exist two vital residues: a lysine that performs the protonation of the ketimine intermediate before release of the product; a histidine that performs most of deprotonation and reprotonation steps in catalytic reaction. There are also three conserved residues (a serine, a lysine and a histidine), which play a very important role in anchoring the phosphate-group of PLP. Sequence alignment for NifS from *A. ferrooxidans* and other sources (Fig.5) also shows that the five residues of Thr73, His100, Ser200, His202, and Lys203 in afNifS are all highly conserved in other species.

The binding sites search results for the modeled structure reveal that three cavity sites are in the afNifS (Fig.4). Interestingly, the conserved residues mentioned above are all located in the site 2. Furthermore, the afNifS and 1P3W are well conserved in both sequences and structures, and their biological functions should be similar. Thus the PLP binds can be presumed in a manner similar to both afNifS and 1P3W. Combined with the search results, the conserved residues in the family of NifS and the experimental results of other analogs to afNifS, site 2 is chosen as the more favorable binding site to dock both the cofactor PLP and substrate cysteine. This site is composed of 18 residues (Asn9, Ala10, Cys71, Gly72, Thr73, Asp76, His100, Ala102, Met148, Asn152, Asp177, Val179, Gln180, Ser198, Ser200, His202, Lys203 and Lys208).

3.3 Interaction between apoenzyme afNifS and cofactor PLP

PLP is an important cofactor in a wide range of biochemical reactions, including amino acid metabolism and antibiotic biosynthesis. All organisms must either produce PLP or acquire it through their diet. The 3D structure of PLP is made up of a pyridine ring connected by a —CHO group, a —OH group, a —CH₃ group and a

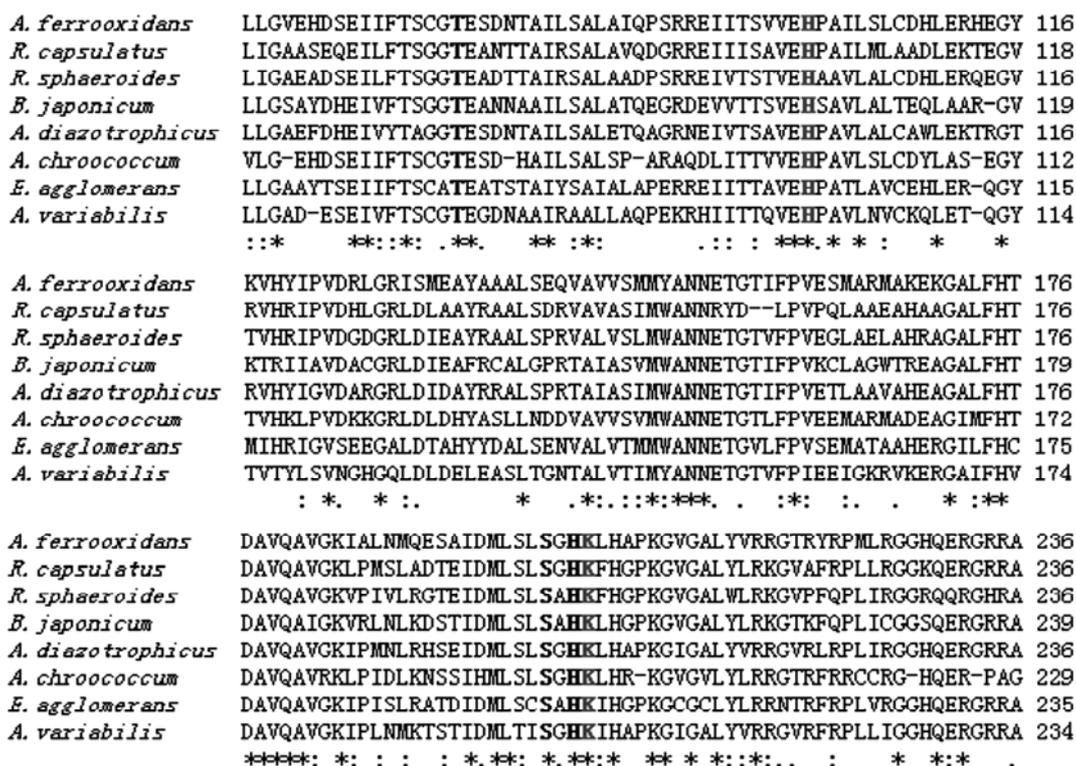


Fig.5 Sequence alignment of PLP-binding domain in NifSs from *A. ferrooxidans* and other sources (Residues conserved in all sequences are marked with “*”; Residues not conserved in all sequences but conserved in some sequences are marked with “.” or “:” based on degree of conservation)

—CH₂PO₄ group.

The binding 3D conformation of the PLP-afNifS docking complex shows that PLP is anchored in the center of the binding pocket and six hydrogen bonds are formed between afNifS and PLP (Fig.6 and Table 1). Three hydrogen bonds are formed between phosphate oxygen atoms of PLP and the side-chains of Thr73, Ser200 and His202 in the same subunit. Additional hydrogen bond interactions occur between the phenolate oxygen of PLP and Gln180, between the aldimine oxygen of PLP and Lys203, and between pyridine N1 of PLP and Asp177. As well known, hydrogen bonds play

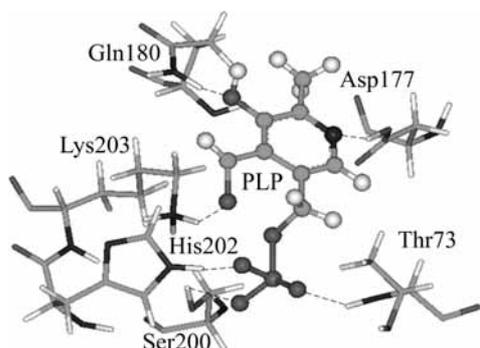


Fig.6 Hydrogen bond interaction between cofactor PLP and afNifS (Hydrogen bonds are represented by dashes)

Table 1 Hydrogen bonds between PLP and residues of afNifS

| afNifS | | PLP | Distance/ | Angle/ |
|---------|------|---------------|-----------|--------|
| Residue | Atom | atom | nm | (°) |
| Lys203 | HZ2 | Aldimine O4A2 | 0.130 | 128.38 |
| Asp177 | HD2 | Pyridine N1 | 0.167 | 156.47 |
| Gln180 | HE21 | Phenolate O3 | 0.156 | 141.38 |
| Thr73 | HN | Phosphate O1P | 0.239 | 156.58 |
| Ser200 | HO | Phosphate O2P | 0.207 | 152.43 |
| His202 | HE2 | Phosphate O3P | 0.145 | 163.46 |

an important role in structure and function of the biological molecules, especially in enzyme catalysis and ligand binding. These hydrogen bonding interactions enhance the stability of the ligand-enzyme complex.

To determine the key residues that comprise the binding pocket of the model, the interaction energies of PLP with each individual amino acid in the enzyme were calculated. The relative importance of each residue in afNifS can be inferred by its rank order of interaction energy. This kind of identification, compared with a definition based on the distance from the substrate, can clearly show the relative significance for every residue.

Table 2 lists the interaction energies for PLP with some residues in afNifS. It is evident from Table 2 that the PLP has favorable interaction energy with

apoenzyme afNifS. The electrostatic energy, the van der Waals energy and the total energy are $-63.756\ 0$, $-67.985\ 4$ and $-131.741\ 4$ kJ/mol, respectively. These results indicate that the attractive interaction is important and in this case both the van der Waals and electrostatic energies are important for determining the binding orientations.

Table 2 Total energy(E_{total}), electrostatic energy(E_{ele}), and van-der-Waals energy(E_{vdw}) between cofactor PLP and individual residues($|E_{\text{total}}| > 0.740\ 0$ kJ/mol listed in energy rank order)

| Residue | $E_{\text{ele}}/(\text{kJ}\cdot\text{mol}^{-1})$ | $E_{\text{vdw}}/(\text{kJ}\cdot\text{mol}^{-1})$ | $E_{\text{total}}/(\text{kJ}\cdot\text{mol}^{-1})$ |
|---------|--|--|--|
| Total | -63.756 0 | -67.985 4 | -131.741 4 |
| Lys203 | -31.683 2 | -7.418 8 | -39.102 0 |
| His100 | -3.465 0 | -16.245 6 | -19.710 6 |
| Asn152 | -9.450 0 | -3.742 2 | -13.192 2 |
| Val179 | -0.273 0 | -6.539 4 | -6.812 4 |
| Asp177 | -4.279 8 | -2.049 6 | -6.329 4 |
| His202 | -1.797 6 | -4.431 0 | -6.228 6 |
| Thr73 | -2.784 6 | -2.734 2 | -5.518 8 |
| Ser200 | -2.381 4 | -2.473 8 | -4.855 2 |
| Gln180 | -0.436 8 | -4.489 8 | -4.926 6 |
| Ala102 | 0.966 0 | -5.602 8 | -4.636 8 |
| Met148 | 0.537 6 | -4.800 6 | -4.263 0 |
| Asp76 | -2.242 8 | -1.138 2 | -3.381 0 |
| Gly72 | 0.226 8 | -3.402 0 | -3.175 2 |
| Glu99 | -2.662 8 | -0.344 4 | -3.007 2 |
| Cys71 | -0.646 8 | -1.251 6 | -1.898 4 |
| Asn151 | -0.163 8 | -0.835 8 | -0.999 6 |
| Ala150 | -0.466 2 | -0.277 2 | -0.743 4 |

The total interaction energy, E_{total} , between the PLP and Lys203 is $-39.102\ 0$ kJ/mol, in which the primary interaction energy is electronic interaction one ($E_{\text{ele}} = -31.683\ 2$ kJ/mol). The residues of Asn152 and Asp177 have the similar behavior as Lys203 and the interaction energies of these residues with PLP are mainly contributed by electronic interaction. But, for the residue of His100, Val179, His202, Gln180, Ala102 and Met148, the interaction energy with PLP is mainly contributed by van der Waals interaction. For the residues of Thr73 and Ser200, the interaction energies with PLP are equal in the gross by van der Waals and electronic interaction.

Through the interaction analysis, the total interaction energies for the above residues are prominent compared with the other residues, so these residues can be conjectured to be important anchoring residues for PLP and have main contribution to the cofactor interaction. Previous experimental results have demonstrated that the NifS-like protein from *E. coli* has seven vital residues[12]: Lys206 that anchors the PLP

group in the active site by formation of an internal aldimine Schiff base; His104 that functions as an acid-base catalyst in several protonation or deprotonation steps during catalysis; Thr76, Ser203 and His205 that play a very important role in anchoring the phosphate-group of PLP; and Asp180 and Gln183 that play an important role in fixing the ring of PLP. A NifS-like protein from *Thermotoga maritima*[13] also has residue counterparts to the NifS from *E. coli*. The residues homologous to Lys206, His104, Thr76, Ser203, His205, Asp180 and Gln183 of NifS-like from *E. coli* are Lys203, His100, Thr73, Ser200, His202, Asp177 and Gln180 of afNifS, respectively(fig.5). To our surprise, the results of the seven residues are well consistent with experimental facts based on the total interacting energy and hydrogen bonds fixation. In addition, four new important residues (Asn152, Val179, Ala102 and Met148) are found according to the interaction energy, which are fitly conserved in NifS from all kinds of sources (Fig.5) but have not been identified before.

Additionally, before the formation of the Michaelis complex between holoenzyme afNifS and substrate cysteine, the PLP is covalently bonded to a lysine [12–15]. The above docking results also support the conclusion because the $-\text{CHO}$ group in PLP is very close to the residue Lys203 (Fig.6) and the interaction energy between them is very large (Table 1).

3.4 Interaction between holoenzyme afNifS and substrate Cysteine

Cysteine is one of the 20 necessary amino acids for protein biosynthesis in all organisms. During the initial steps of substrate reaction, the cysteine enters the active pocket and forms the Michaelis complex; the substrate amino group is deprotonated by His100; and then nucleophilically attack on C4A atom in PLP to form the external aldimine (Fig.7(a)). The substrate docking in this work mainly aimed at the formation of Michaelis complex. The results indicate that the amino group in cysteine is near to His100 and PLP and the distances between the HN2 atom in amino group of cysteine and the NE2 atom in His100, and between the HN3 atom in amino group of cysteine and the C4A atom in PLP are 0.172 nm and 0.235 nm (Fig.7(b) and Fig.8), respectively, which are well agreed with the initial process of substrate reaction mechanism (Fig.7(a)).

Table 3 lists the interaction energies of cysteine with afNifS holoenzyme. The substrate cysteine has favorable interaction energy with holoenzyme afNifS. The electrostatic energy, the van der Waals energy and the total interaction energy for the total residues are $-45.015\ 6$, $-35.040\ 6$ and $-80.056\ 2$ kJ/mol, respectively. This interaction energy corresponds only to the enthalpic contribution to the free energy of binding. The results

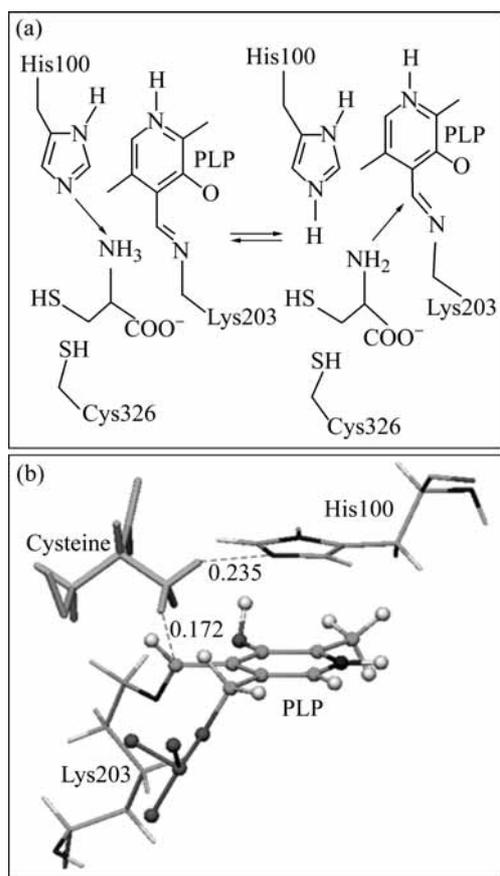


Fig.7 Interaction between substrate cysteine and holoenzyme afNifS: (a) Schematic of initial steps for possible reaction mechanism of afNifS; (b) Position of substrate docking (PLP is represented by stick-and-ball; His100 and Lys203 are represented by thin sticks; and cysteine is represented by sticks)

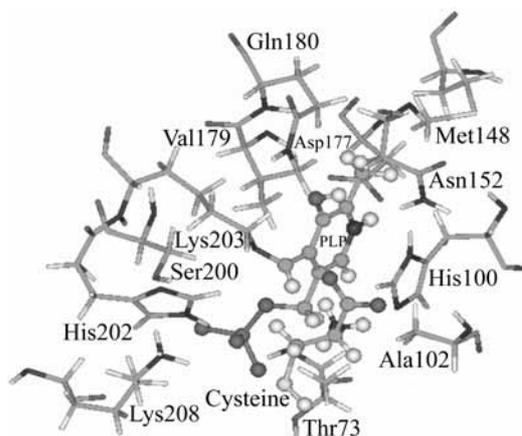


Fig.8 Key residues around cofactor PLP and substrate cysteine (PLP and cysteine in center is represented by sticks and balls; all key residues around it are represented by sticks)

indicate that in this case both the van der Waals and electrostatic energies are important for determining the substrate binding orientation, and the large attractive interaction energy is helpful for effectively recruiting cysteine. The total energy, E_{total} , between the cysteine

Table 3 Total energy(E_{total}), electrostatic energy(E_{ele}), and van-der-Waals energy(E_{vdw}) between substrate cysteine and individual residues($|E_{\text{total}}| > 1.680 0 \text{ kJ/mol}$ listed in energy rank order)

| Residue | $E_{\text{ele}}/(\text{kJ}\cdot\text{mol}^{-1})$ | $E_{\text{vdw}}/(\text{kJ}\cdot\text{mol}^{-1})$ | $E_{\text{total}}/(\text{kJ}\cdot\text{mol}^{-1})$ |
|---------|--|--|--|
| Total | -45.015 6 | -35.040 6 | -80.056 2 |
| His100 | -22.041 6 | -10.290 0 | -32.331 6 |
| PLP | -11.503 8 | -7.001 4 | -18.505 2 |
| Lys203 | -6.388 2 | -2.507 4 | -8.895 6 |
| Lys208 | -4.834 2 | -0.604 8 | -5.439 0 |
| Ala102 | -3.754 8 | -0.205 8 | -3.960 6 |
| Arg352 | -1.797 6 | -0.252 0 | -2.049 6 |
| Asp177 | 1.268 4 | -3.087 0 | -1.818 6 |
| Asp76 | 1.289 4 | -3.045 0 | -1.755 6 |
| Glu153 | 1.289 4 | -2.973 6 | -1.684 2 |

and His100 is $-32.331 6 \text{ kJ/mol}$, in which the primary interaction energy is electronic energy ($E_{\text{ele}} = -22.041 6 \text{ kJ/mol}$). The residues of Lys203, Lys208, Ala102 and cofactor PLP have the similar behavior as His100 and the interaction energies of these residues with cysteine are mainly contributed by electronic interaction. Interestingly, the total interaction of His100, Lys203 and cofactor PLP are prominent compared with the other residues, which indicates that they are important active components for cysteine and have main contribution to the substrate interaction. Obviously, the results of the cofactor PLP and the two residues of His100 and Lys203 are well consistent with experimental facts of NifS from other species[13]. However, Lys208 and Ala102 are new identified important residues according to the interaction energies, which are also fitly conserved in NifS from all kinds of sources (Fig.5) but has not been identified before.

All of the key residues identified in the PLP docking and in the cysteine docking are shown in Fig.8. They are composed of a pocket embraced PLP and cysteine, and form a site characterized by an unusually low rate of mutation. Such spatial site is sure to identify evolutionarily privileged functional site. Obviously in this work, this site is the active site candidate directly for catalytic function and cofactor binding.

4 Conclusions

1) The reliable integral 3D molecular model of the enzyme encoded by the gene *iscS-2* from *A. ferrooxidans* was built. From the information of modeled structure and interactions with PLP and cysteine, this gene can be confirmed to encode the NifS and subsequently play a sulfur traffic role in nitrogenase maturation of this bacterium.

2) According to the docking with cofactor PLP, the PLP is firmly anchored in the binding pocket of NifS from *A. ferrooxidans*. The residues of Lys203, His100, Thr73, Ser200, His202, Asp177 and Gln180 are directly responsible for PLP binding and catalytic function in NifS from *A. ferrooxidans*.

3) According to the docking with substrate cysteine, the cysteine can be effectively recruited into the active site of NifS from *A. ferrooxidans* and subsequently activated. The residues of Lys208, His106 and Ala102 are of vital importance in the catalysis of cysteine.

4) The detailed 3D structure, interaction information and identified key residues are helpful for guiding site-directed mutagenesis investigation, understanding the catalytic mechanism of the enzyme and subsequently insight into the nitrogen fixation metabolism of *A. ferrooxidans* so as to finally serve for industrial bioleaching.

References

- [1] RAWLINGS D E, KUSANO T. Molecular genetics of *Thiobacillus ferrooxidans* [J]. Microbiological Reviews, 1994, 58(1): 39–55.
- [2] BAKER B J, BANFIELD J F. Microbial communities in acid mine drainage [J]. FEMS Microbiology Ecology, 2003, 44(2): 139–152.
- [3] TUOVINEN O H, NIEMELA S I, GYLLENBERG H G. Tolerance of *Thiobacillus ferrooxidans* to some metals [J]. Antonie Van Leeuwenhoek, 1971, 37(4): 489–496.
- [4] BANERJEE P C. Genetics of metal resistance in acidophilic prokaryotes of acidic mine environments [J]. Indian Journal of Experimental Biology, 2004, 42(1): 9–25.
- [5] QUATRINI R, APPIA-AYME C, DENIS Y, RATOUCHEIAK J, VELOSO F, VALDES J, LEFIMIL C, SILVER S, ROBERTO F, ORELLANA O, DENIZOT F, JEDLICKI E, HOLMES D, BONNEFOY V. Insights into the iron and sulfur energetic metabolism of *Acidithiobacillus ferrooxidans* by microarray transcriptome profiling [J]. Hydrometallurgy, 2006, 83(1/4): 263–272.
- [6] DEAN D R, BOLIN J T, ZHENG L. Nitrogenase metalloclusters: Structures, organization, and synthesis [J]. Journal of Bacteriology, 1993, 175(21): 6737–6744.
- [7] PETERS J W, SZILAGYI R K. Exploring new frontiers of nitrogenase structure and mechanism [J]. Current Opinion in Chemical Biology, 2006, 10(2): 101–108.
- [8] JACOBSON M R, CASH V L, WEISS M C, LAIRD N F, NEWTON W E, DEAN D R. Biochemical and genetic analysis of the *nifUSVWZM* cluster from *Azotobacter vinelandii* [J]. Molecular and General Genetics, 1989, 219(1/2): 49–57.
- [9] JOHNSON D C, DEAN D R, SMITH A D, JOHNSON M K. Structure, function, and formation of biological iron-sulfur clusters [J]. Annual Review of Biochemistry, 2005, 74: 247–281.
- [10] MIHARA H, ESAKI N. Bacterial cysteine desulfurases: Their function and mechanisms [J]. Applied Microbiology and Biotechnology, 2002, 60(1/2): 12–23.
- [11] ZHENG L, WHITE R H, CASH V L, DEAN D R. Mechanism for the desulfurization of L-cysteine catalyzed by the *IscS* gene product [J]. Biochemistry, 1994, 33(15): 4714–4720.
- [12] CUPP-VICKERY J R, URBINA H, VICKERY L E. Crystal structure of *IscS*, a cysteine desulfurase from *Escherichia coli* [J]. Journal of Molecular Biology, 2003, 330(5): 1049–1059.
- [13] KAISER J T, CLAUSEN T, BOURENKOW G P, BARTUNIK H D, STEINBACHER S, HUBER R. Crystal structure of a NifS-like protein from *Thermotoga maritima*: Implications for iron sulfur cluster assembly [J]. Journal of Molecular Biology, 2000, 297(2): 451–464.
- [14] FUJII T, MAEDA M, MIHARA H, KURIHARA T, ESAKI N, HATA Y. Structure of a NifS homologue: X-ray structure analysis of CsdB, an *Escherichia coli* counterpart of mammalian selenocysteine lyase [J]. Biochemistry, 2000, 39(6): 1263–1273.
- [15] TIRUPATI B, VEY J L, DRENNAN C L, BOLLINGER J M. Kinetic and structural characterization of Slr0077/SufS, the essential cysteine desulfurase from *Synechocystis* sp. PCC 6803 [J]. Biochemistry, 2004, 43(38): 12210–12219.
- [16] EISENHABER F, PERSSON B, ARGOS P. Protein structure prediction: Recognition of primary, secondary, and tertiary structural features from amino acid sequence [J]. Critical Reviews in Biochemistry and Molecular Biology, 1995, 30(1): 1–94.
- [17] BLUNDELL T L, SIBANDA B L, STERNBERG M J E, THORNTON J M. Knowledge-based prediction of protein structures and the design of novel molecules [J]. Nature, 1987, 326: 347–352.
- [18] LUTHY R, BOWIE J U, EISENBERG D. Assessment of protein models with three-dimensional profiles [J]. Nature, 1992, 356(6364): 83–85.
- [19] MORRIS A L, MACARTHUR M W, HUTCHINSON E G, THORNTON J M. Stereochemical quality of protein structure coordinates [J]. Proteins, 1992, 12(4): 345–364.

(Edited by YANG Bing)