

*nifH*基因的生物信息学分析

第15小组

报告人：王庆峰

成员：刘晓梅 王秀呈 孟利娟

背景介绍

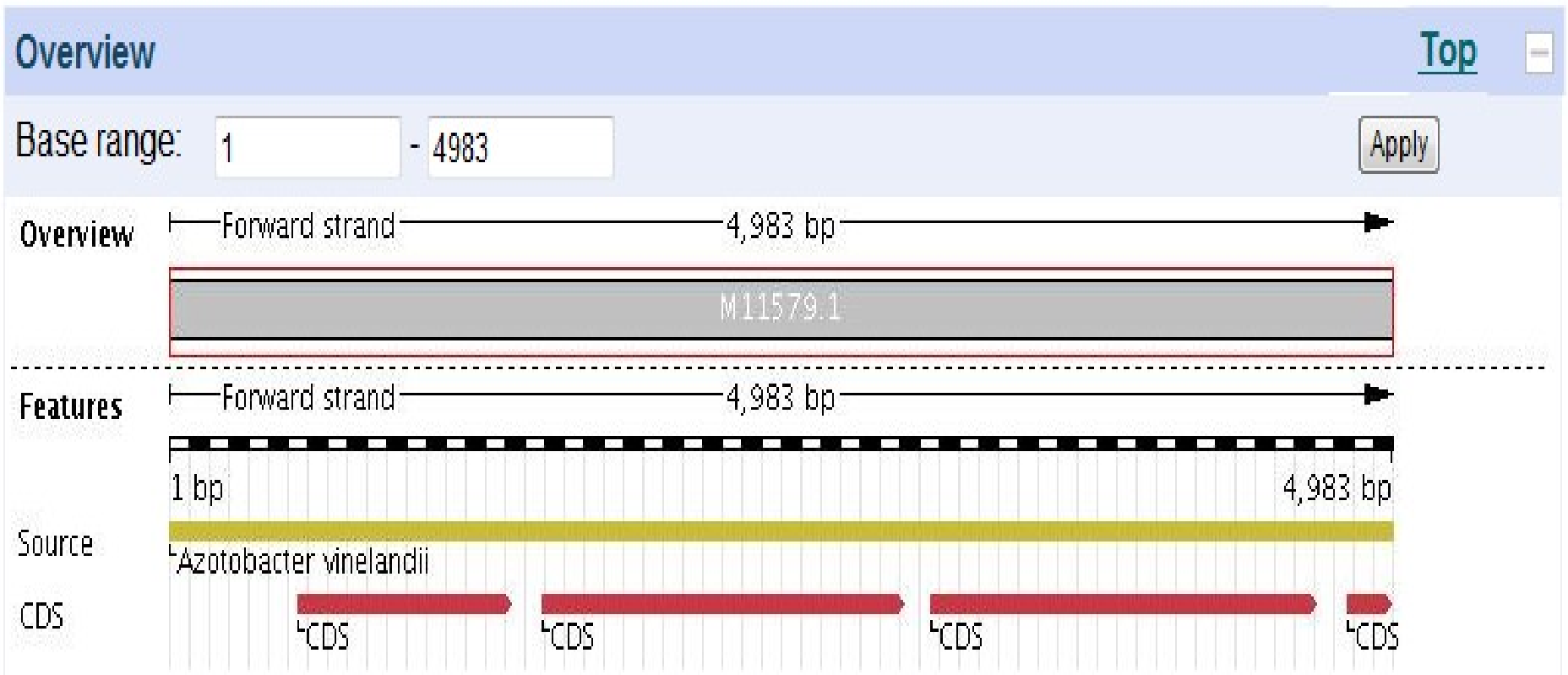
人口、环境、资源已成为当今社会愈演愈烈的三大问题，粮食资源虽然较之前有所缓解但仍然不能满足当今社会的需求。无机化肥N,P,K的施用虽然在一定程度上提高了粮食的产量，但同时也带来了严重的环境问题，例如土壤板结，水体污染等。因此，生物肥料成为当今科学家们研究的主要方向。

在很久之前就发现大豆等豆科作物根部存在根瘤菌。这类细菌能够直接利用环境中的氮气，并将其转化成植物能够利用的氮元素。

研究发现，在这类固氮微生物中，行使固氮功能的为一类*nif*基因。

*nifH*为这类固氮基因中的一员,*nifH*的基因序列保守性不强,不同菌株之间*nifH*的结构也不同。因此可以利用*nifH*基因来研究土壤中固氮菌类群的多样性,即利用*nifH*基因来进行物种鉴定。本课题主要就是用*nifH*来进行物种的鉴定。

在EBI中搜索 *nifH* 的编码序列信息



Other Features

[Top](#)

Base range:

- Show main features only

CDS

525..1397

codon_start

1

transl_table

11

note

Fe protein

translation

MAMRQCAIYGKGGIGKSTTTQNLVAALAEMGKKVMIVGCDPKADSTRILILHKAQNTIMEMAAEAGTVEDLELEDVLKAG
YGGVKCVESGGPEPGVGCAGRGVITAINFLEEGAYEDDLDFVFYDVLGDVVCVCGGFAMPPIRENKPQEIYIVCSGEMMAMY
AANNISKGIVKYANSGSVRLGGLICNSRNTDREDELI IALANKLGTQMIHFVPRDNVVQRAEIRRMTVIEYDPKAKQADE
YRALARKVVDNKLKLLVIPNPITMDELEELLMEFGIMEVEDESIVGKTAAEV

↓ Coding:

[AAA22142](#)

→ GOA

[P00459](#)

→ InterPro

[IPR000392](#), [IPR005977](#)

→ PDB

[1DE0](#), [1FP6](#), [1G1M](#), [1G20](#), [1G21](#), [1G5P](#), [1M1Y](#), [1M34](#), [1N2C](#), [1NIP](#), [1RW4](#), [1XCP](#), [1XD8](#),
[1XD9](#), [1XDB](#), [2AFH](#), [2AFI](#), [2AFK](#), [2C8V](#), [2NIP](#)→ UniProtKB/Swiss-
Prot[P00459](#)

蛋白质的性质

Protein attributes

Sequence length	290 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level

注释

General annotation (Comments)

Function	The key enzymatic reactions in nitrogen fixation are catalyzed by the nitrogenase complex, which has 2 components: the iron protein (component 2) and a component 1 which is either a molybdenum-iron protein, a vanadium-iron, or an iron-iron protein. HAMAP-Rule MF_00533
Catalytic activity	$8 \text{ reduced ferredoxin} + 8 \text{ H}^+ + \text{N}_2 + 16 \text{ ATP} + 16 \text{ H}_2\text{O} = 8 \text{ oxidized ferredoxin} + \text{H}_2 + 2 \text{ NH}_3 + 16 \text{ ADP} + 16 \text{ phosphate}$. HAMAP-Rule MF_00533
Cofactor	Binds 1 4Fe-4S cluster per dimer.
Subunit structure	Homodimer.
Post-translational modification	The reversible ADP-ribosylation of Arg-101 inactivates the nitrogenase reductase and regulates nitrogenase activity By similarity . HAMAP-Rule MF_00533
Miscellaneous	This subunit is associated with the molybdenum-iron nitrogenase component 2. HAMAP-Rule MF_00533
Sequence similarities	Belongs to the NifH/BchL/ChlL family .

蛋白质的主要结合位点

Keywords

Biological process	Nitrogen fixation
Ligand	4Fe-4S ATP-binding Iron Iron-sulfur Metal-binding Nucleotide-binding
Molecular function	Oxidoreductase
PTM	ADP-ribosylation
Technical term	3D-structure Direct protein sequencing

基因本体论

Gene Ontology (GO)

Biological_process	nitrogen fixation Inferred from electronic annotation. Source: UniProtKB-KW
Cellular_component	molybdenum-iron nitrogenase complex Inferred from electronic annotation. Source: InterPro
Molecular_function	4 iron, 4 sulfur cluster binding Inferred from electronic annotation. Source: UniProtKB-KW ATP binding Inferred from electronic annotation. Source: UniProtKB-KW carbonyl sulfide nitrogenase activity Inferred from electronic annotation. Source: EC metal ion binding Inferred from electronic annotation. Source: UniProtKB-KW nitrogenase activity Inferred from electronic annotation. Source: EC

蛋白序列的性质

Sequence annotation (Features)

Feature key	Position (s)	Length	Description	Graphical view	
Molecule processing					
<input type="checkbox"/>	Initiator methionine	1	1	Removed Ref.3	
<input checked="" type="checkbox"/>	Chain	2 – 290	289	Nitrogenase iron protein 1 HAMAP-Rule MF_00533	
Regions					
<input type="checkbox"/>	Nucleotide binding	10 – 17	8	ATP HAMAP-Rule MF_00533	
Sites					
<input type="checkbox"/>	Metal binding	98	1	Iron-sulfur (4Fe-4S); shared with dimeric partner	
<input type="checkbox"/>	Metal binding	133	1	Iron-sulfur (4Fe-4S); shared with dimeric partner	
Amino acid modifications					
<input type="checkbox"/>	Modified residue	101	1	ADP-ribosylarginine; by dinitrogenase reductase ADP-ribosyltransferase By similarity	
Experimental info					
<input type="checkbox"/>	Mutagenesis	16	1	K → Q or P: Loss of nitrogen fixation. Ref.5	
<input type="checkbox"/>	Sequence conflict	145	1	A → P in AAA22142. Ref.2	

蛋白质的二级结构

Secondary structure



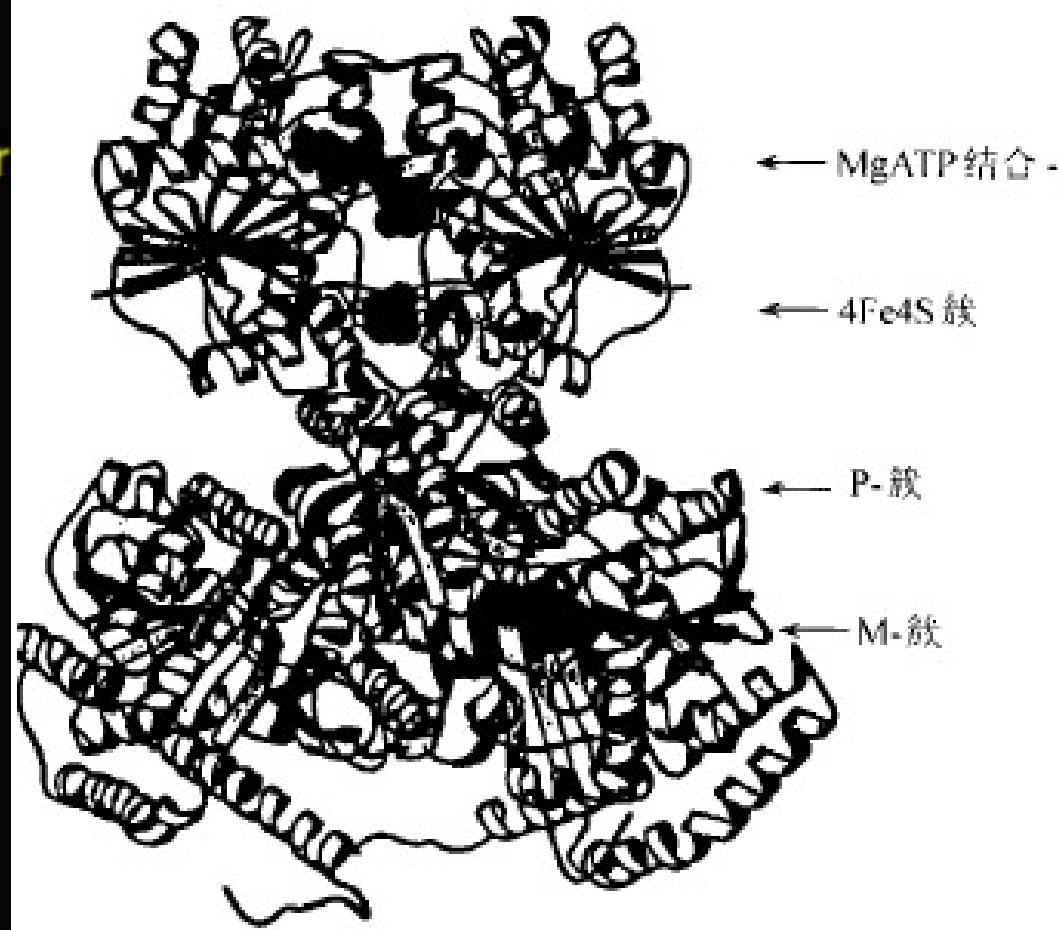
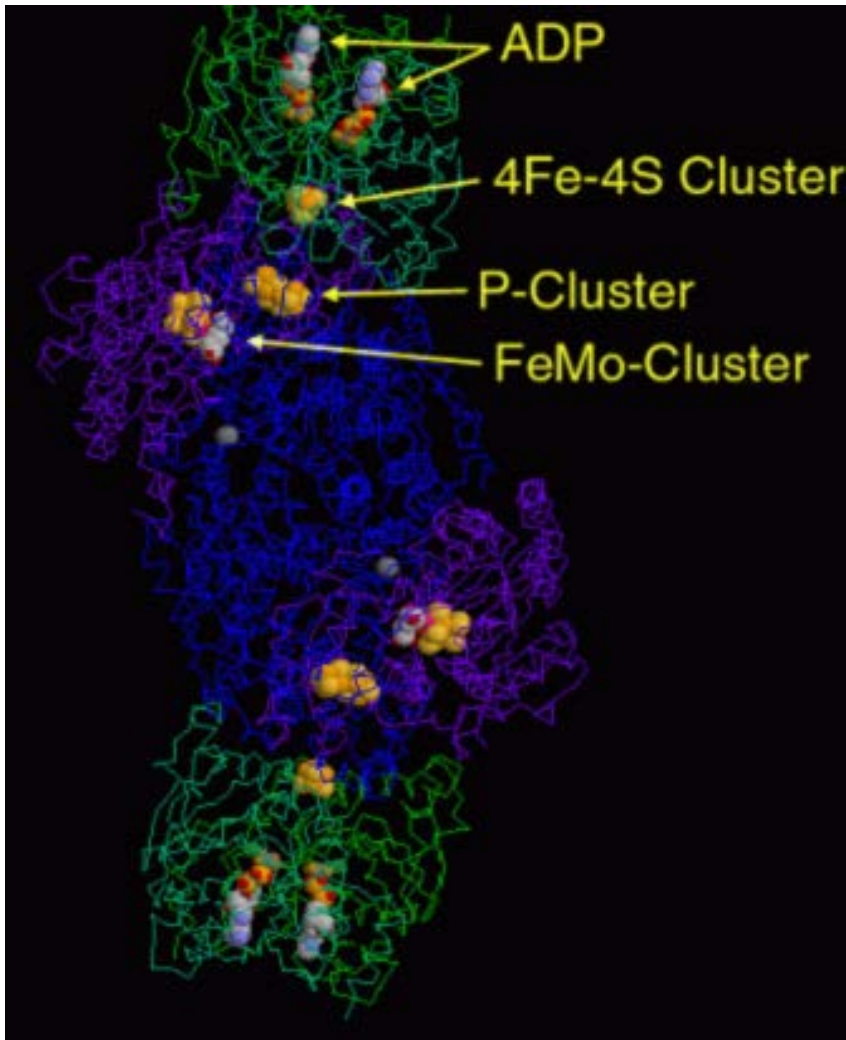


图4 铁蛋白-钼铁蛋白的镶嵌模式^[10,13]

PDB id: **1de0**

[Links](#)

Name: **Oxidoreductase**

Title: Modulating the midpoint potential of the [4Fe-4S] cluster of the nitrogenase Fe protein

Structure: Nitrogenase iron protein. Chain: a, b. Mutation: yes

Source: *Azotobacter vinelandii*. Organism_taxid: 354

Biol. unit: Dimer (from [PQS](#))

Resolution: 2.40Å R-factor: 0.221 R-free: 0.281

Authors: S.B.Jang,L.C.Seefeldt,J.W.Peters

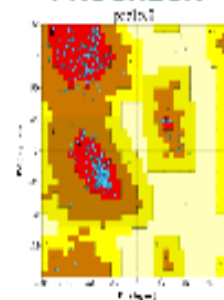
Key ref: S.B.Jang et al. (2000). Modulating the midpoint potential of the [4Fe-4S] cluster of the nitrogenase Fe protein. *Biochemistry*, 39, 641-648.



PubMed id: [10651628](#) DOI: [10.1021/bi991694v](#)

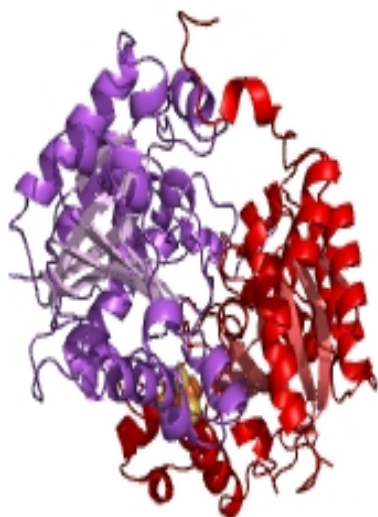
Date: 12-Nov-99 Release date: 09-Feb-00

PROCHECK



Headers

References



Jmol *Strep*

Protein chains **A** **B**



[P00459](#) (NIFH1_AZOVI) - Nitrogenase iron protein 1



Seq: 290 a.a.

Struc: 289 a.a.*

Key: Family PfamA domain Secondary structure CATH domain

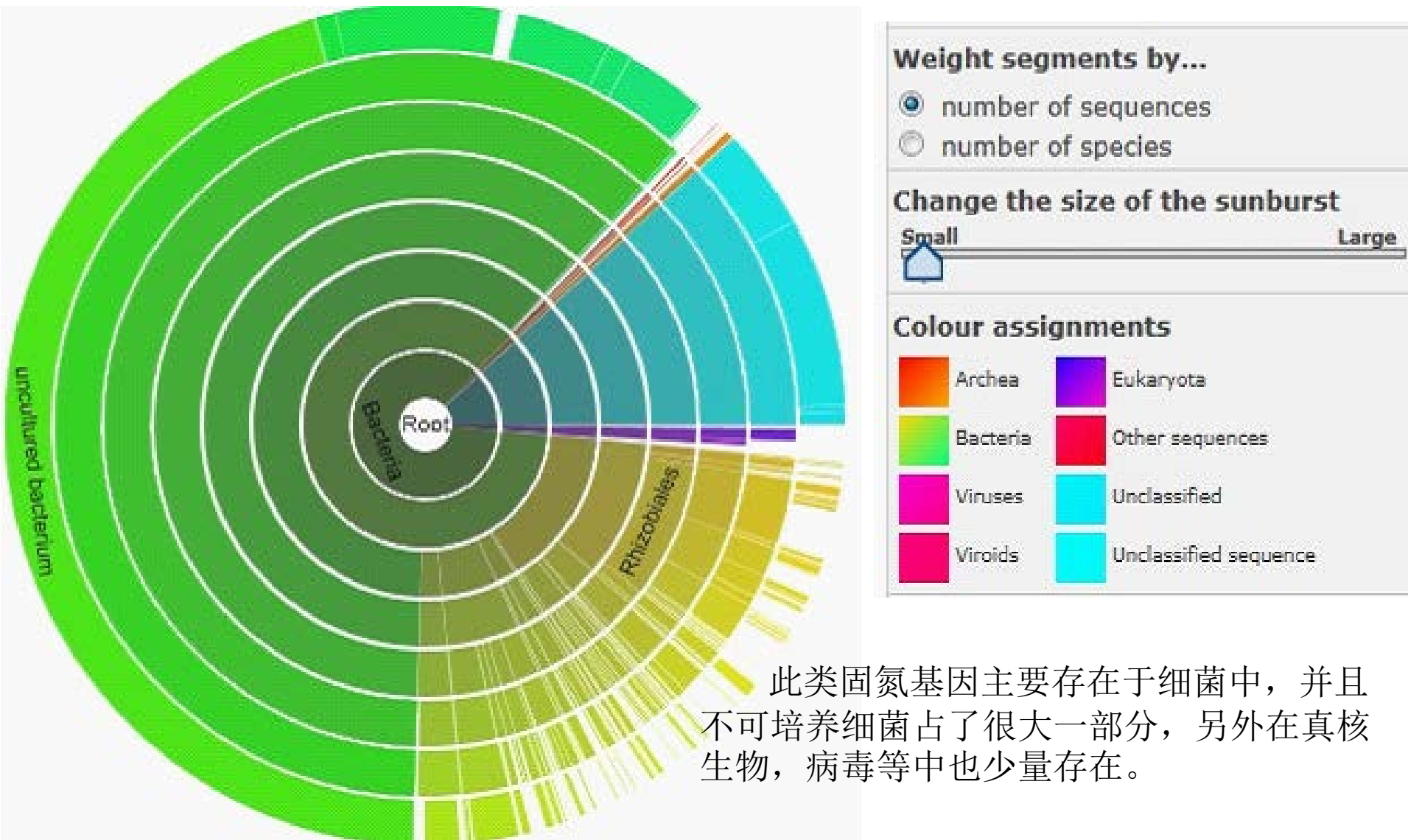
* PDB and UniProt seqs differ at 1 residue position (black cross)

Contents

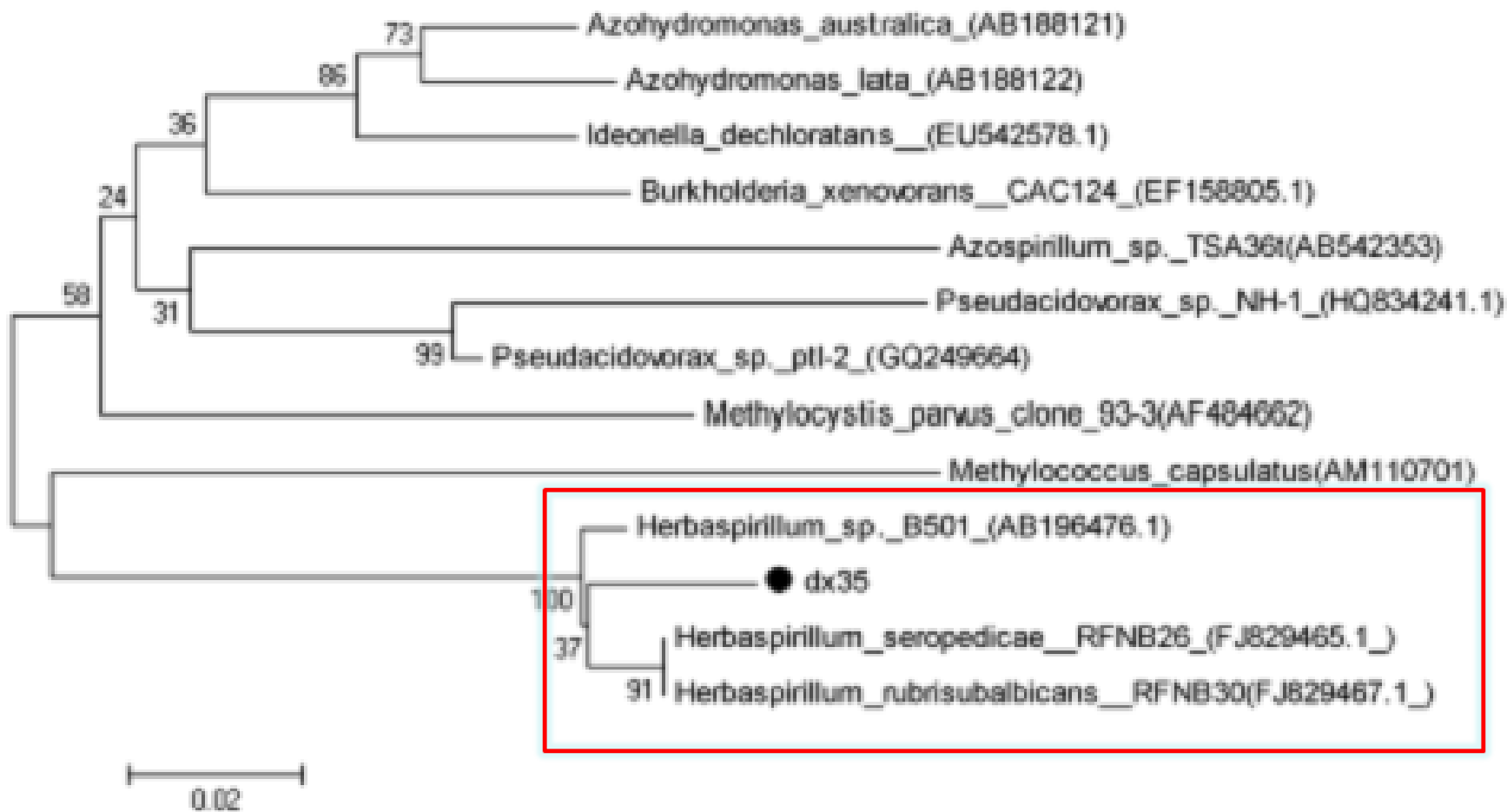
Protein chains

A **B** 289 a.a.*

固氮基因分布及系统发育分析



此类固氮基因主要存在于细菌中，并且不可培养细菌占了很大一部分，另外在真核生物，病毒等中也少量存在。



DX35在NCBI中进行比对，挑选出序列相似性在97%以上的不同属种的代表菌种，构建系统发育树，得到的结果如上图所示，DX35与*Herbaspirillum*(草螺菌属)聚在了一起。

固氮酶的结构

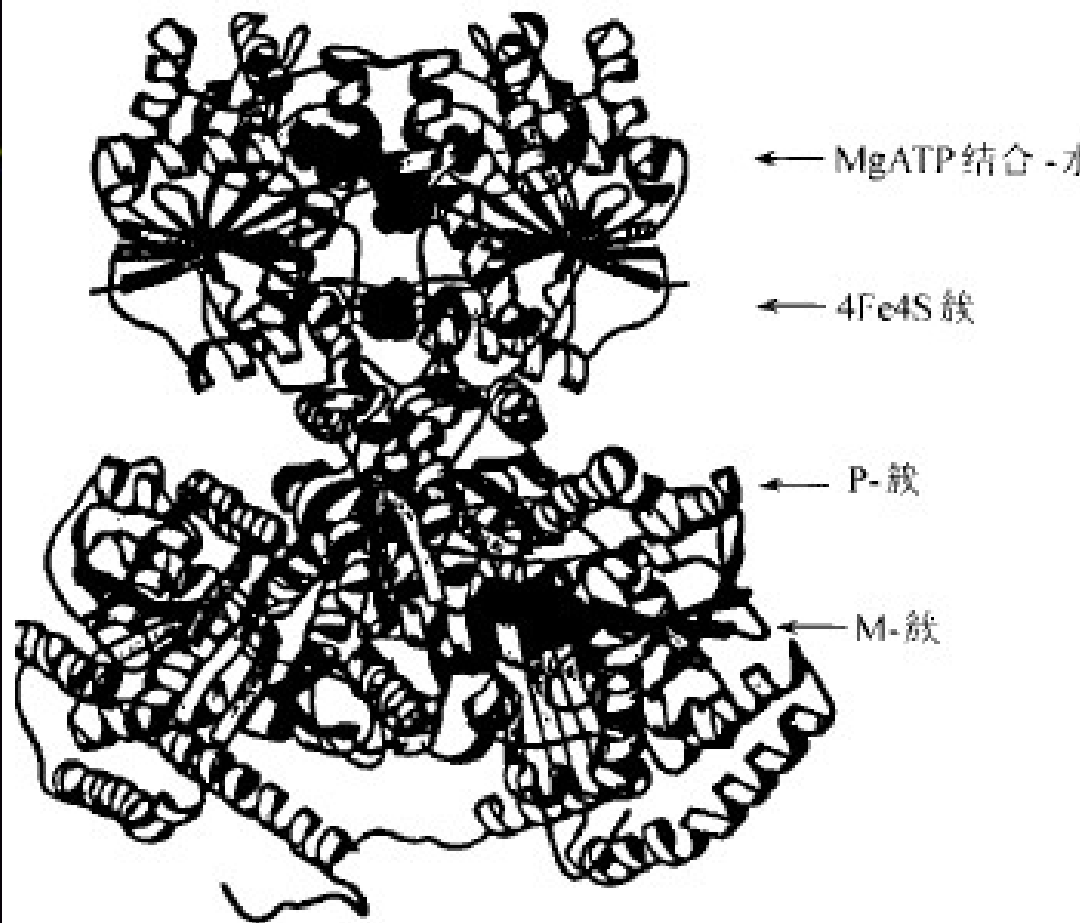
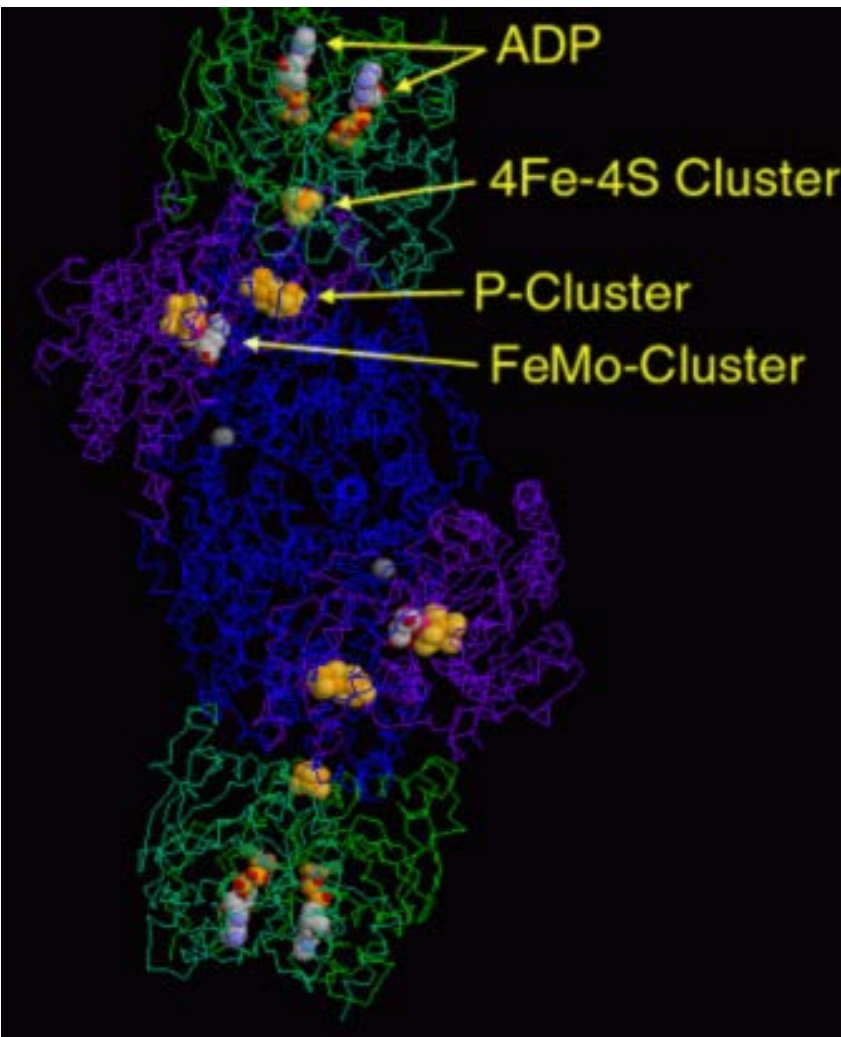
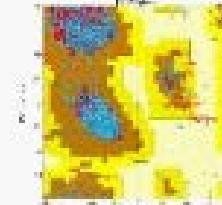
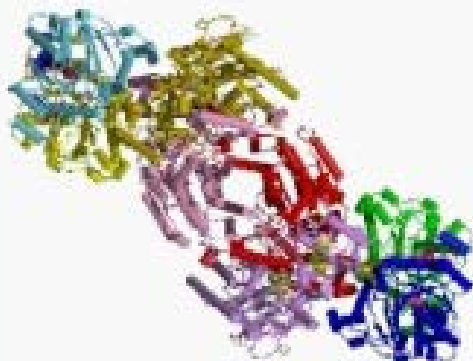


图4 铁蛋白-钼铁蛋白的镶嵌模式^[10,13]

生物固氮主要是固氮酶发挥作用。固氮酶是同型二聚体，是双组分酶复合体包括二氮酶（钼铁蛋白）和二氮还原酶（铁蛋白）。钼铁蛋白是 $\alpha_2\beta_2$ 型四聚体， α 亚基由*nifD*编码， β 亚基由*nifK*编码。二氮酶还原酶是由两个相同亚基组成的 γ_2 型二聚体，由*nifH*编码。

PDB id: **1g21**[Links](#)

PROCHECK



[Headers](#)
[References](#)

Name: **Oxidoreductase**

Title: Mgatp-bound and nucleotide-free structures of a nitrogenase protein complex between leu127del-fe protein and the mofe protein

Structure: Nitrogenase molybdenum-iron protein alpha chain. Chain: a, c. Synonym: nitrogenase component i, dinitrogenase, nifd. Nitrogenase molybdenum-iron protein beta chain. Chain: b, d. Synonym: nitrogenase component i, dinitrogenase, nifk. Nitrogenase iron protein. Chain: e, f, g, h. Synonym: nitrogenase component ii, nitrogenase reductase.

Source: Azotobacter vinelandii. Organism_taxid: 354. Organism_taxid: 354

Biol. unit: Octamer (from [PQS](#))

Resolution: 3.00Å R-factor: 0.238 R-free: 0.285

Authors: H.-J. Chiu, J. W. Peters, W. N. Lanzetta, M. J. Ryle, L. C. Seefeldt, J. B. Howard, D. C. Rees

Key ref: H. Chiu et al. (2001). MgATP-Bound and nucleotide-free structures of a nitrogenase protein complex between the Leu 127 Delta-Fe-protein and the MoFe-protein. *Biochemistry*, 40, 641-650. PubMed id: [11170380](#) DOI: [10.1021/bi001645e](#)

Date: 16-Oct-00 Release date: 31-Jan-01

Jmol

Contents

Protein chains

- [A](#) [C](#) 476 a.a. ?
- [B](#) [D](#) 522 a.a. ?
- [E](#) [F](#) [G](#) [H](#) 268 a.a. ?

Ligands

- [SF4](#) ×2
- [ATP](#) ×4
- [HCA](#) ×2
- [CFM](#) ×2
- [CLF](#) ×2

Metals

- [_MG](#) ×4
- [_CA](#) ×2

* Residue conservation analysis

Protein chains [A](#) [C](#)

[View](#) [Zoom](#) [Rotate](#) [ES-ES](#) [Pfam](#) [Info](#)

[P07328](#) ([NIFD_AZOV1](#)) - Nitrogenase molybdenum-iron protein alpha chain

Seq: 492 a.a.
Struc: 476 a.a.

Protein chains [B](#) [D](#)

[View](#) [Zoom](#) [Rotate](#) [ES-ES](#) [Pfam](#) [Info](#)

[P07329](#) ([NIFK_AZOV1](#)) - Nitrogenase molybdenum-iron protein beta chain

Seq: 523 a.a.
Struc: 522 a.a.

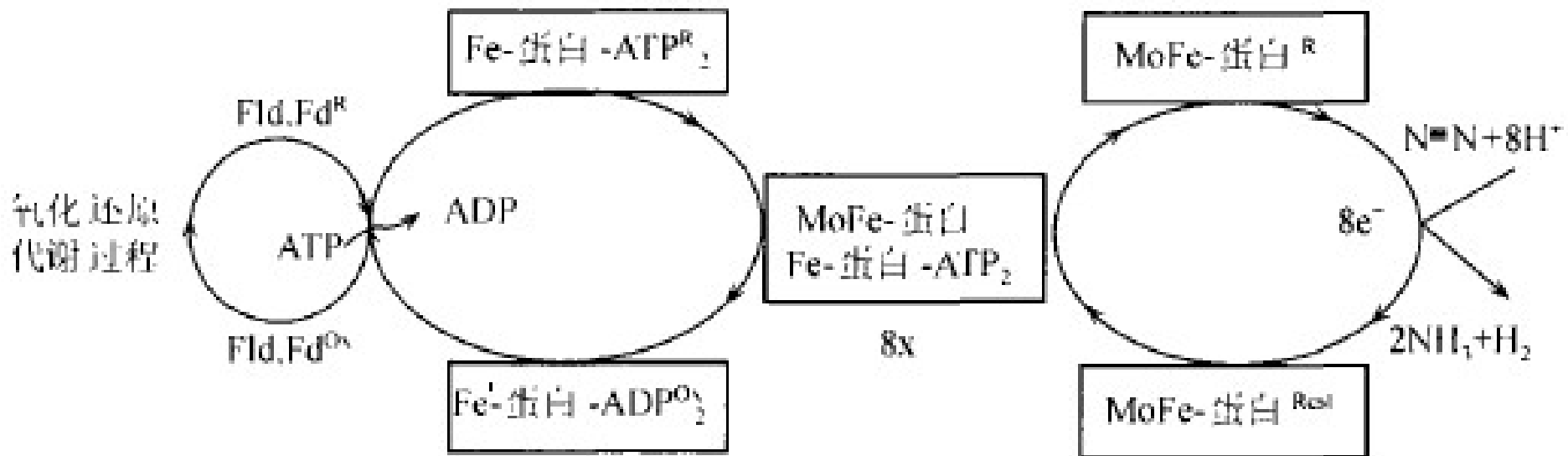
Protein chains [E](#) [F](#) [G](#) [H](#)

[View](#) [Zoom](#) [Rotate](#) [ES-ES](#) [Pfam](#) [Info](#)

[P00459](#) ([NIFH1_AZOV1](#)) - Nitrogenase iron protein 1

Seq: 290 a.a.
Struc: 268 a.a.

Key: PfamA domain PfamB domain Secondary structure CATH domain



ATP与铁蛋白结合, 铁蛋白-ATP 与钼铁蛋白结合成复合物,ATP水解,复合物内发生电子转移(在生理条件下由铁氧还蛋白或黄素蛋白提供的电子, 经铁蛋白的 $4\text{Fe}4\text{S}$ 转移到钼铁蛋白的 $8\text{Fe}7\text{S}$, 继而进入 Fe Moco 。), 底物结合与还原, 铁蛋白- ADP 复合物与钼铁蛋白分离, ATP 替换出 ADP 又形成铁蛋白 ATP 复合物, 如此完成一次循环。



不同物种铁蛋白氨基酸序列比对

2nipa	(1)	amrgCAIYGkggIgKstTTQNLVAALAEmgkkVMIVGCdpkadSTrlllh
1cp2a	(1)	mrqVAIYGkggIgKstTTQNLISGLhamgkIMVVGCDpkadSTrlllg
		bbbbbbb aaaaaaaaaaaaaa bbbbb aaaa
2nipa	(51)	skagntImema aa agtvedleledVlkagyggVkcVESGGpepgvgcaGr
1cp2a	(50)	GlaQksVldTlreeg--edVeldsIlkegyggIrCveSggpepgvgcaGr
		aaaaa aaaa bb 333 bbbb
2nipa	(101)	GVitAInFLeegAyeddLdFVFYDVLGdvVcgGFaMPIrenkAqeIYIV
1cp2a	(98)	gIitSInmLeqlgAytdLdYVFYDVLGdvvcgGFaMPIregkAqeIYIV
		aaaaaaaaa bbbbbbb aaaa bbbbb
2nipa	(151)	CSGemAMyaAnnISkgIvkyansgsVrLGGLICnSrntrEdeLIiaLA
1cp2a	(148)	ASGemMALyaAnnISkgIqkyAksggVrLGGIICNsrkvaneyeLLdaFA
		b aaaaaaaaaaaaaa bbbbbbb aaaaaa
2nipa	(201)	nkLgTgmIhfvprdnVvgrAeirrmTVIeydpkakQadeYraLArkVvdn
1cp2a	(198)	keLgSgLIhfvprspmvtkAeinkqTVIeydptceQAeeYreLArkVdan
		aaa bbbbb aaaaaa aaaa aaaaaaaaaaaaaa
2nipa	(251)	kllviPnpItmdeLeeLLmefgimevedesivgkta
1cp2a	(248)	elfviPkpMtqerLeeIlmqyg
		aaaaaaaaa

不同物种之间序列存在很大的差异性，因此可以以这类基因为标记研究固氮微生物的多样性。

谢谢