Journal of Integrative Plant Biology 2008, **50** (7): 784–796

Invited Review

Perspectives in Biological Nitrogen Fixation Research

Qi Cheng*

(*Department of Biochemistry, Redox Biology Center, University of Nebraska-Lincoln*, Lincoln, Nebraska 68588, USA)

Abstract

Nitrogen fixation, along with photosynthesis is the basis of all life on earth. Current understanding suggests that no plant fixes its own nitrogen. Some plants (mainly legumes) fix nitrogen via symbiotic anaerobic microorganisms (mainly rhizobia). The nature of biological nitrogen fixation is that the dinitrogenase catalyzes the reaction-splitting triple-bond inert atmospheric nitrogen (N_2) into organic ammonia molecule (NH_3). All known nitrogenases are found to be prokaryotic, multi-complex and normally oxygen liable. Not surprisingly, the engineering of autonomous nitrogen-fixing plants would be a long-term effort because it requires the assembly of a complex enzyme and provision of anaerobic conditions. However, in the light of evolving protein catalysts, the anaerobic enzyme has almost certainly been replaced in many reactions by the more efficient and irreversible aerobic version that uses O_2. On the other hand, nature has shown numerous examples of evolutionary convergence where an enzyme catalyzing a highly specific, O_2-requiring reaction has an oxygen-independent counterpart, able to carry out the same reaction under anoxic conditions. In this review, I attempt to take the reader on a simplified journey from conventional nitrogenase complex to a possible simplified version of a yet to be discovered light-utilizing nitrogenase.

Key words: aerobic; anaerobic; anoxic; convergent enzyme; light-independent; light-utilizing; nitrogen fixation; nitrogenase; oxic; protochlorophyllide reductase.

Cheng Q (2008). Perspectives in biological nitrogen fixation research. *J. Integr. Plant Biol.* **50**(7), 784–796.

Available online at www.jipb.net

Nitrogen was discovered by Daniel Rutherford in 1772. It was found to be so inert that Antoine Lavoisier named it "azote", meaning "without life". Dinitrogen (N_2) has a triple bond and does not readily accept or donate electrons. As a gas or liquid, nitrogen is colorless and odorless. Two allotropic forms of solid nitrogen exist, with the transition from the α to the β form taking place at -237 °C. Like the global metabolism of many elements, the N cycle can be summarized as transfer among the inorganic forms of nitrogen. Certain microorganisms have the ability to use the renewable source of energy to fix atmospheric nitrogen (constituting 78% of air) under mild conditions, such as normal temperature and normal pressure. Nitrogen fixation is a key process in which molecular nitrogen is reduced to form ammonia, which is the form of nitrogen that is used by living systems for the synthesis of many bioorganic compounds. Biologically-fixed nitrogen could be directly "absorbed" by plants and keep the environment almost "untouched". Crop rotation with legumes has been recognized to increase soil fertility and agricultural productivity since ancient China and Rome. However, the science behind such practice was not revealed until Boussingault experimented with leguminous crops fixing N_2 in 1838; Hellriegel and Wilfarth showed definitive evidence for N_2 fixation by microbes in legumes in 1886. The Haber-Bosch industrial process was established in 1906, which uses a catalytic agent (iron with a small amount of aluminum added) at high pressure (as much as 5.06×10^7 Pa) and high temperature (600−800 K) which normally consumes fossil fuel. Annually, approximately 2.5×10^{11} kg NH_3 are fixed from the atmosphere by biological nitrogen fixation (by legumes and cyanobacteria) and approximately 8×10^{10} kg NH_3 are manufactured by ammonia industry. Lightning worldwide may also contribute approximately 1×10^{10} kg NH_3/year. Currently, approximately 2 tons of industrially-fixed nitrogen are needed as fertilizer for crop production to equal the effects of 1 ton of nitrogen biologically-fixed by legume crops. Therefore, biologically-fixed nitrogen influences the global nitrogen cycle substantially less than industrially-fixed nitrogen. One day, this situation needs to be changed.

On the other hand, world population has now been increasingly relying on nitrogen fertilizers in order to keep up with the demands of food and economic growth rates. As one can see,

Received 29 Mar. 2008 Accepted 16 Apr. 2008

*Author for correspondence.

Tel: +1 402 403 0567;

Fax: +1 402 472 7842;

E-mail: <qcheng2@unl.edu>.

doi: 10.1111/j.1744-7909.2008.00700.x

a large proportion of fertilizers come from ammonia industry. However, less than 30% of synthetic fertilizers would actually be utilized, the unused chemicals sprayed on crops would be lost in the field and could subsequently cause serious environmental problems, let alone industrial pollution. Biological nitrogen fixation has the advantage of being environmental friendly and therefore would be ideal for sustainable agriculture. Research in this field has pivotal importance and would be significantly beneficial. Enormous progress in almost all aspects of biological nitrogen fixation has been made in the past century, especially in the recent two decades, in genetics and biochemistry, culminating in the determination of the crystallographic structures of both nitrogenase components. However, yet more dynamic studies need to be carried out by biochemists, chemists, biophysicists, crystallographers, theoreticians and geneticists in order to completely understand the nature of the process and make possible use of it.

Assembly of the Nitrogenase Complex: Biochemistry and Genetics

The most well-studied nitrogenase contains two metallo-components, dinitrogenase (molybdenum-iron (MoFe) protein) and dinitrogenase reductase (Fe protein). The overall stoichiometry of dinitrogen reduction by nitrogenase (EC 1.18.2.1) is: $N_2 + 8H^+ + 8e^- + 16MgATP \rightarrow 2NH_3 + H_2 + 16MgADP + 16Pi$.

Nitrogenase turnover requires an electron donor in addition to adenosine triphosphate (ATP). Electrons are generated *in vivo* either oxidatively or photosynthetically, depending on the organism. These electrons are transferred to flavodoxin or ferredoxin, a (4Fe-4S)-containing electron carrier that transfers an electron to the Fe protein of nitrogenase, beginning a series of oxido-reduction cycles. Two molecules of MgATP bind to the reduced Fe protein and are hydrolyzed to drive an electron from the Fe protein to the MoFe protein. The actual reduction of N_2 occurs on the MoFe protein in a multistep reaction. Electron transfer must occur six times per each fixed N_2 molecule so that a total of 12 ATPs are required to fix one N_2 molecule. However, nitrogenase also reduces protons to H_2 a reaction which consumes two electrons. Therefore, the total cost of N_2 reduction is eight electrons transferred and 16 MgATPs hydrolyzed (Figure 1).

Adenosine triphosphate hydrolysis, electron transfer and substrate reduction are the key steps for nitrogenase turnover. The breakthrough of nitrogen fixation research culminated with the structural characterization of the nitrogenase components in late 20th century (Georgiadis et al. 1992; Kim and Rees 1992, 1993; Schindelin et al. 1997), which contributed significantly to our understanding of enzymatic nitrogen reduction. The locations of the metal clusters in the nitrogenase components persuasively suggest a general sequence of electron transfer: (4Fe4S) → (8Fe8/7S) → (7Fe9S[Mo/V/Fe]) → substrate

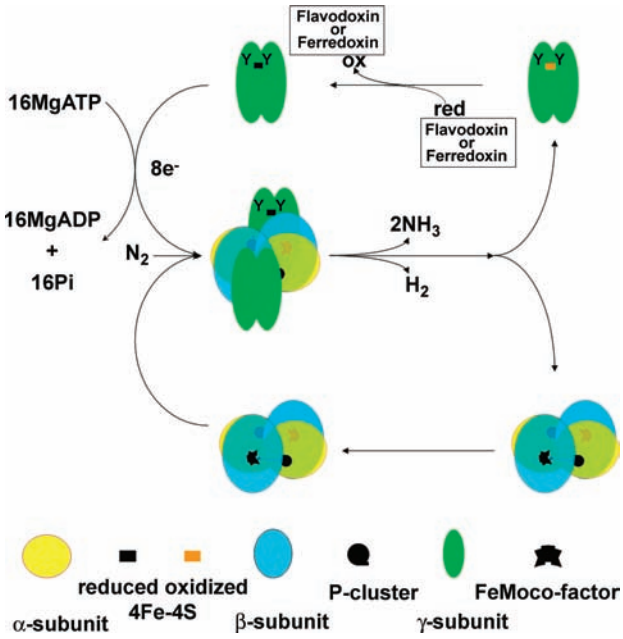

Figure 1. The mechanism of nitrogenase turnover.

The reduction of N_2 occurs on the MoFe protein ($\alpha 2\beta 2$ hetero-tetramer) in a multiple-step reaction with the Fe protein ($\gamma 2$ homodimer). Electron transfer six times per N_2 molecule fixed and nitrogenase also reduces protons to H_2, consuming two electrons. The total cost of N_2 reduction is therefore 16 MgATPs hydrolyzed and eight electrons transferred. ATP, adenosine triphosphate.

(Figure 2). The P-cluster (8Fe8/7S) is the electron "pool" which has almost equal distance (~14 Å) to either Fe protein (4Fe4S) or iron-molybdenum cofactor (FeMoco). The MoFe protein is an $\alpha 2\beta 2$ hetero-tetramer providing substrate-binding and -reduction sites. The tetramer contains 30 Fe and two Mo atoms, which are distributed between two types of cluster, the P-cluster (8Fe8S) in the Rees model or (8Fe7S) in the Bolin model and the FeMo-cofactor (seven Fe, one Mo, nine S, one homocitrate) (Rees et al. 1993; Bolin et al. 1993). The α and β subunits are composed of 491 and 522 amino acids, respectively, in the case of Av1, with a total molecular weight of approximately 240 kDa. The tetramer interface is dominated by interactions between helices from the two β subunits, along with a cation binding site, presumably occupied by calcium that is coordinated by residues from both β subunits. The contacts between the $\alpha\beta$ pairs are almost entirely between the β subunits. The P-cluster, which may function in electron transfer between the (4Fe-4S) cluster and the FeMoco, is located approximately 10 Å beneath the protein surface, on the twofold axis that approximately relates the α and β subunits. The FeMoco, which may function in substrate binding and 6 reduction, is also buried approximately 10 Å beneath the protein surface in an environment primarily provided by the α subunit. These amino acid environments not

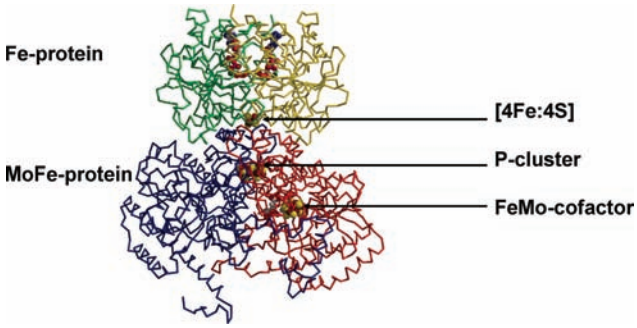

Figure 2. Complex of the nitrogenase proteins (Fe protein and MoFe protein).

The individual subunits of each Fe protein colored green and yellow. The MoFe α- and β-subunits are colored red and blue, respectively. Non-protein groups are shown in a space-filling representation, with fluorine and magnesium colored orange and green, respectively. Transduction pathway coupling the nucleotide and cofactor sites in the nitrogenase complex: (4Fe4S)-cluster, P-cluster, and FeMo-cofactor sites (Howard and Rees 2006).

only stabilize the protein structure but also influence substrate reduction properties, such as the rate, substrate specificity and the products.

In the nitrogenase cycle, the role for ATP hydrolysis is to control the electron-transfer "gate" between protein components. How this is accomplished is still one of the two main unanswered questions about the nitrogenase mechanism (the other one being how substrates are reduced at the FeMoco). There is excellent evidence that the FeMo-cofactor clusters act as the enzyme's substrate-binding and -reducing site, but exactly how and where substrates bind and are activated remains controversial. A very close observation was reported with a high-resolution structure of the MoFe protein of nitrogenase, which reveals a previously unrecognized interstitial atom in the FeMo cofactor that may possibly be nitrogen (Einsle et al. 2002). Electron transfer from the Fe protein to the MoFe protein is coupled to the hydrolysis of MgATP which is followed by dissociation of the protein-protein complex. The electron transfer step is an essential function of the Fe protein, because the MoFe protein alone will not reduce N_2 in the absence of the Fe protein, despite the fact that the MoFe protein component (cofactor) can be reduced by other electron donors.

In nitrogen-fixing bacteria, nitrogenase is encoded by a set of operons which includes regulatory genes (such as *nifLA*), structural genes (such as *nifHDK*) and other supplementary genes. The free-living diazotrophic bacterium, *Klebsiella pneumoniae*, has been the most extensively analyzed and provides a model for studies of nitrogenase regulation, synthesis and assembly. A 24 kb base pair DNA region, contains the entire *K. pneumoniae nif* cluster, which includes 20 genes (Figure 3). *nifHDK* are the three structural genes encoding for the three subunits of Mo-nitrogenase. In most nitrogen fixing prokaryotes, these three genes form one transcriptional unit, with a promoter in front of the *nifH* gene. The maturation of apo-Fe protein (NifH) requires the products of *nifH*, *nifM*, *nifU* and *nifS*, while that of apo-MoFe protein requires at least six genes *nifE*, *nifN*, *nifV*, *nifH*, *nifQ* and *nifB* which are required for the biosynthesis of FeMoco. There is considerable homology between *nifDK* and *nifEN*, and it has been speculated that the *nifEN* products might form a scaffold for FeMoco biosynthesis that later shifts FeMoco to the *nifDK* complex. The *nifB* gene product, termed NifB-co, is an iron- and sulfur-containing precursor of FeMoco. The *nifQ* gene product might be involved in the formation of a molybdenum-sulfur precursor to FeMoco and it has a typical motif characteristic of metal-binding sequences (Cys-X4-Cys-X2-Cys-X5-Cys). The *nifV* gene encodes a homocitrate synthase and is required for the synthesis of FeMoco. *nifW* is not required for the initial assembly of the MoFe protein but rather may be necessary to protect the MoFe protein from O_2 damage. The *nifY* gene product has a function similar to γ protein. *nifF* and *nifJ* encode components of a specific electron transfer pathway in which electrons are donated from pyruvate to a flavodoxin and hence to the Fe protein of nitrogenase. Although *nifM* has been found necessary for the maturation and stabilization of the *nifH* product, its actual role has never been defined. In contrast, *nifS* and *nifU* are ubiquitous in various organisms (Dos Santos et al. 2004).

Nitrogenase Has "Siblings": Double-Bond Reductases for Chlorophyll Biosynthesis in Photosynthesis

Chlorophyll is essential for life in the biosphere, playing an important role in the energy absorption and transduction processes of photosynthetic organisms. Chlorophyll catalyzes the conversion of solar energy to chemical energy via the process of photosynthesis. Approximately 250–300 of them transfer the absorbed light energy through neighboring pigments to the "special pair" of chlorophylls in a reaction center. This special pair of chlorophylls in photosystems I and II are the primary electron donors that drive the conversion of light into chemical energy to be conserved in nicotinamide adenine dinucleotide phosphate (NADPH) and ATP. The most important pigment molecule in photosynthesis is chlorophyll *a*, which absorbs light only at certain wavelengths. The process of photosynthesis can capture more energy if it uses other molecules, accessory pigments, to absorb the energy from other wavelengths and pass it on to chlorophyll *a*. The reduction of NADPH::protochlorophyllide (Pchlide) is a key step in this biosynthesis of chlorophyllide (Chl) during the greening of phototrophic organisms. The process is catalyzed by the key enzyme NADPH::protochlorophyllide oxidoreductase (light-operative protochlorophyllide oxidoreductase (LPOR); EC 1.3.1.33). With the exception of angiosperms,

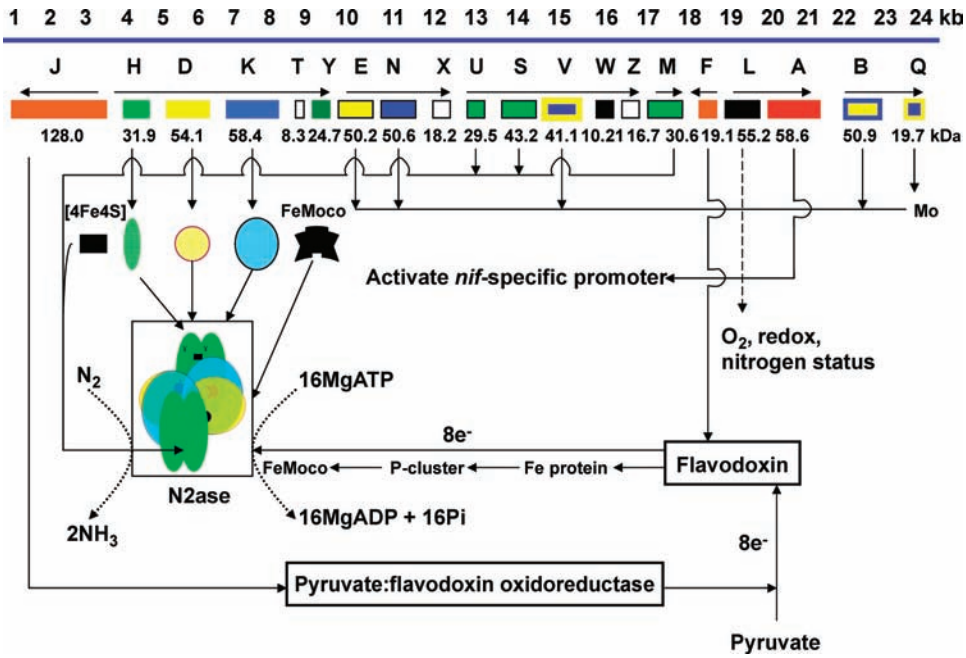

Figure 3. *Klebsiella pneumoniae nif* cluster including 20 genes (~24 kb).

Transcriptional orientation, protein product molecular weights and known functions are indicated.
ADP, adenosine diphosphate; ATP, adenosine triphosphate.

a light-independent chlorophyll biosynthesis pathway has been found and catalyzed by the light-independent enzyme dark-operative protochlorophyllide oxidoreductase (DPOR) at the key step of converting protochlorophyllide to chlorophyllide a (Figure 4).

Genes for such light-independent pathways of chlorophyll biosynthesis were first discovered when the complete chloroplast genome sequence from the liverwort, *Marchantia polymorpha*, was determined (Ohyama et al. 1986). A gene designated *frxC* (later renamed *chlL*) at that time was found to have a high degree of similarity with *nifH* of nitrogenase. Genetic studies in *Rhodobacter capsulatus* also demonstrated that three genes, *bchL*, *bchN* and *bchB*, were involved in protochlorophyllide reduction (Zsebo and Hearst 1984; Bauer et al. 1988; Yang and Bauer 1990). Sequence analysis revealed the surprising finding that these genes, together with genes identified from cyanobacterium, algae and gymnosperms, termed *chlL*, *chlN* and *chlB*, exhibit striking similarities to the three subunits of nitrogenase (Lidholm and Gustafsson 1991; Fujita et al. 1992, 1993; Li et al. 1993; Liu et al. 1993). This suggests that light-independent protochlorophyllide reductase and nitrogenase share a common evolutionary ancestor (Burke et al. 1993; Fujita 1996). In nitrogenase, the Fe protein (structurally encoded by *nifH*) specifically transfers electrons to the MoFe protein (structurally encoded by *nifD* and *nifK*) in a reaction that is coupled with the hydrolysis of MgATP (see the nitrogenase mechanism). ChlL is similar to NifH (~35%) (Figure 5). ChlNB is similar to both NifD and NifK (~19%) (Fujita et al. 1991, 1993) and, in particular, has cystein residues at positions equivalent to Cys α62, α88, α154, β70, β52, β95 and β153, which coordinate the P clusters in nitrogenase MoFe protein. In addition, some similarity (~16%) between ChlN and ChlB is also apparent, as it is between NifD and NifK (Fujita 1996). Similarities between the subunits of DPOR and those of nitrogenase suggest that DPOR has a molecular architecture similar to that of nitrogenase (Figure 6). It has become increasingly evident that, in the DPOR complex, ChlL contains a (4Fe-4S) cluster, functioning as a specific donor of electrons to the other component, consisting of ChlN and ChlB, which catalyze the reduction of Pchlide (Fujita and Bauer 2000; Cheng et al. 2005a; Brocker et al. 2008). In the synthesis of Bchl, the reduction of the chlorin B-ring double-bond in chlorophyllide *a* (λ_{max}: 661 nm) to bacteriochlorophyllide *a* (λ_{max}: 771 nm) is sequentially catalyzed by chlorophyllide reductase (COR) (Nomata et al. 2006). COR is encoded by *bchX*, *bchY* and *bchZ*, which are also homologous to *chlL*, *chlN* and *chlB*, respectively (Burke-Aquero 1992; Burke et al. 1993). The similarities among DPOR, COR and nitrogenase suggest an evolutionary relationship. It appears that nitrogenase may have more than one "sibling".

Of great interest, from the protein point of view, is that nitrogenase remains "prokaryotic" while its "sibling" (DPOR) evolves into a higher version (LPOR), ultimately becoming

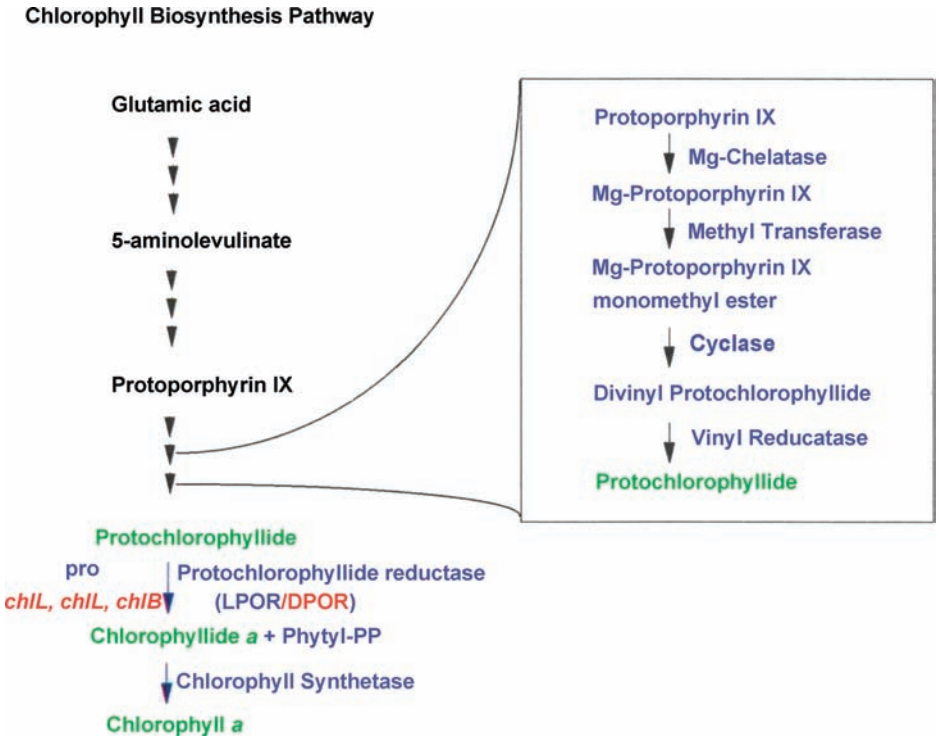

Figure 4. Chlorophyll biosynthesis pathway.

Protochlorophyllide reductase catalyzing protochlorophyllide to chlorophyllide *a* by dark-operative protochlorophyllide oxidoreductase (DPOR; exist up to gymnosperm) or light-operative protochlorophyllide oxidoreductase (LPOR; dominant in angiosperm).

dominant in higher plants (angiosperms). DPOR and LPOR are very different proteins and structurally unrelated, but carry out similar reactions with different chemistry (Figure 7). LPOR occurs in all oxygenic photosynthetic organisms. With the exception of angiosperms, LPOR coexists with DPOR, although DPOR is not essential and may not become redundant. LPOR is apparently absent from anoxygenic bacteria, which have only DPOR enzymes (Suzuki and Bauer 1995; Schoefs and Franck 2003). The LPOR gene originated in the cyanobacterial genome before the divergence of oxic photosynthetic organisms. The photosynthetic eukaryotes obtained their LPOR homologs through endosymbiotic gene transfer (Yang and Cheng 2004). In green algae and higher plants, LPOR is encoded in the nucleus, translated as a precursor polypeptide in the cytosol, and ultimately post-translationally processed and imported into plastids. Comparison of the primary and secondary structure of LPORs with other proteins requiring NADPH as co-factor has revealed that LPOR belongs to an extended superfamily of NAD(P)H-accepting oxidoreductase, termed the short-chain dehydrogenase/reductase (SDR) family (Figure 8). A most striking feature of LPOR is that it requires light for its catalytic activity, one of the only two photo-enzymes (Dahlin et al. 1999).

Efforts to Introduce Nitrogenase Gene into Chloroplast Genome in Non-Legume Plant

Since more than three decades ago when the *nif* genes were successfully transferred into *Escherichia coli* from the nitrogen-fixing bacterium, *K. pneumoniae* (Dixon and Postgate 1972), the transfer of *nif* genes directly into plant cell to create diazotrophic plants has also been considered (Earl and Ausubel 1982). Several possible locations for expression of nitrogenase in the plant cell have been suggested (Merrick and Dixon 1984) and attempts to express *nif* genes in higher plants was achieved by targeting NifH or NifH plus NifM into tobacco chloroplasts, although in both cases NifH was found in the chloroplast stromal fraction at a very low level (Dowson-Day et al. 1991). There are advantages for the chloroplast being a potential environment for nitrogen fixation: the developed plastid carries out photosynthesis and thus provides a major source of ATP. The chloroplast is also a major site for ammonia assimilation because both the GS and the GOGAT pathways exist in the chloroplast. From an engineering point of view, the plastid provides a convenient location for the introduction of *nif* genes because chloroplast genes are expressed in a prokaryotic-like fashion, allowing translation of polycistronic messages. *Chlamydomonas* possesses a

```
                    motif A
NIFH_AZOCH      MAMRQCAIYGKGGIGKSTTTQNLVAALAEMGKKVMIVGCDPKADSTRLILHSKAQNTIME 60
NIFH_AZOVI      MAMRQCAIYGKGGIGKSTTTQNLVAALAEMGKKVMIVGCDPKADSTRLILHSKAQNTIME 59
NIFH_KLEPN      -TMRQCAIYGKGGIGKSTTTQNLVAALAEMGKKVMIVGCDPKADSTRLILHAKAQNTIME 59
CHLL_CHLRE      ---MKLAVYGKGGIGKSTTSCNISIALRKRGKKVLQIGCDPKHDSTFTLTGFLIPTIIDT 57
                    : *:***********: *:  ** : ****: :***** ***  :      . *

NIFH_AZOCH      MAAEAGTVEDLELEDVLKVGYGGVKCVESGGPEPGVGCAGRGVITAINFLEEEGAYEDDL 120
NIFH_AZOVI      MAAEAGTVEDLELEDVLKAGYGGVKCVESGGPEPGVGCAGRGVITAINFLEEEGAYEDDL 119
NIFH_KLEPN      MAAEVGSVEDLELEDVLQIGYGDVRCAESGGPEPGVGCAGRGVITAINFLEEEGAYEDDL 119
CHLL_CHLRE      LSSKDYHYEDIWPEDVIYGGYGGVDCVEAGGPPAGAGCGGYVVGETVKLLKELNAFF-EY 116
                ::::    **:  ***:  ***.* *.*:*** .*.**.*  *  ::::*:* .*:  :

                    motif B
NIFH_AZOCH      DFVFYDVLGDVVCGGFAMPIRENKAQEIYIVCSGEMMAMYAANNISKGIVKYANSGSVRL 180
NIFH_AZOVI      DFVFYDVLGDVVCGGFAMPIRENKAQEIYIVCSGEMMAMYAANNISKGIVKYANSGSVRL 179
NIFH_KLEPN      DFVFYDVLGDVVCGGFAMPIRENKAQEIYIVCSGEMMAMYAANNISKGIVKYAKSGKVRL 179
CHLL_CHLRE      DVILFDVLGDVVCGGFAAPLN--YADYCIIVTDNGFDALFAANRIAASVREKARTHPLRL 174
                *.:::************ *:.   *:   ** .. : *::***.*: .: : *.:  :**

NIFH_AZOCH      GGLICNSRNTDREDELIIALAAKLGTQMIHFVPRDNVVQRAEIRRMTVIEYDPTAKQADE 240
NIFH_AZOVI      GGLICNSRNTDREDELIIALANKLGTQMIHFVPRDNVVQRAEIRRMTVIEYDPKAKQADE 239
NIFH_KLEPN      GGLICNSRQTDREDELIIALAEKLGTQMIHFVPRDNIVQRAEIRRMTVIEYDPACKQANE 239
CHLL_CHLRE      AGLIGN-RTSKRD--LIDKYVEACPMPVLEVLPLIEEIRISRVKGKTLFEMSNKNNMTSA 231
                .*** * * :.*:  **   .      ::..:*  : :: :.::  *::* .   : :.

NIFH_AZOCH      YRTLARKVVENKMLIIPNPITMDELEALLMEFGVMEEEDESIVGKAAAAEE--------- 291
NIFH_AZOVI      YRALARKVVDNKLLVIPNPITMDELEELLMEFGIMEVEDESIVGKTAEEV---------- 289
NIFH_KLEPN      YRTLAQKIVNNTMKVVPTPCTMDELESLLMEFGIMEEEDTSIIGKTAAEENAA------- 292
CHLL_CHLRE      HMDGS-KGDNSTVGVSETPSEDYICNFYLNIADQLLTEPEGVIPRELADKELFTLLSDFY 290
                :   : *  :..: :  .*      :  *   . :  *  .:: :

NIFH_AZOCH      ---
NIFH_AZOVI      ---
NIFH_KLEPN      ---
CHLL_CHLRE      LKI 293
```

Figure 5. Amino acid sequence alignment between ChlL and NifH.

Similar nucleotide binding motif A (GXXXXGK[15]S), motif B (D^{125}XXG) and conserved cysteins for liganding the (4Fe-4S) cluster (Cys^{97} and Cys^{132}) are found. The identical residues are indicated by asterisks. CHL_CHLRE, *Chlamydomonas reinhardtii* ChlL; NIFH_KLEPN, *Klebsiella pneumoniae* NifH; NIFH_AZOCH, *Azotobacter chroococcum* NifH; NIFH_AZOVI, *Azotobacter vinelandii* NifH.

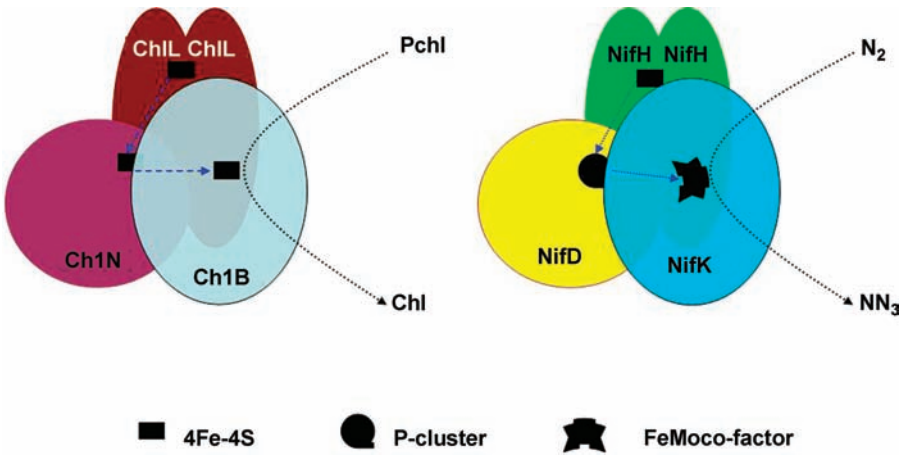

Figure 6. Dark-operative protochlorophyllide oxidoreductase (DPOR) structural model and comparison with nitrogenase.

Electron transfer pathway in DPOR (ChlL [4Fe4S]-cluster → ChlNB [4Fe4S]-center → ChlNB [4Fe4S]-center) versus in nitrogenase (NifH [4Fe4S]-cluster → NifDK [P-cluster] → NifDK [FeMo-cofactor, FeMoco]).

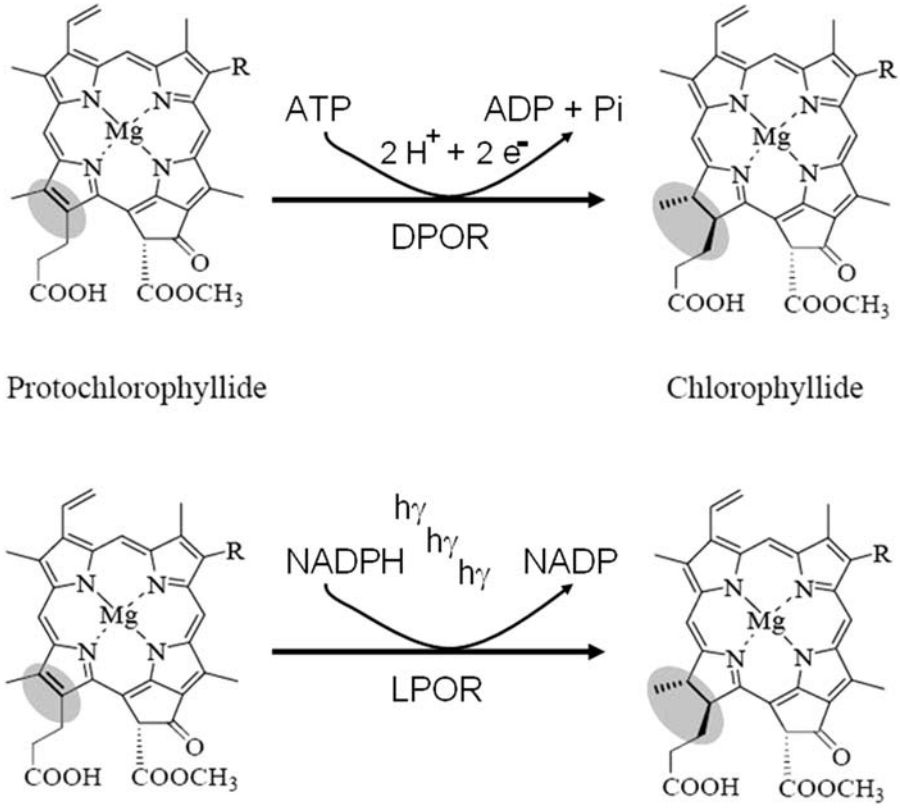

Figure 7. Dark- and light-operative protochlorophyllide oxidoreductase (DPOR/LPOR) are different proteins, but carry out similar reactions with different chemistry.

ADP, adenosine diphosphate; ATP, adenosine triphosphate.

light-independent pathway for chlorophyll biosynthesis, with one of the enzymes in the pathway having potential structural and functional homologies to nitrogenase. The products of the *Chlamydomonas reinhardtii chlL*, *N* and *B* genes are structurally similar to the three subunits of nitrogenase, with the strongest sequence identity between *nifH* and *chlL* (nucleotide sequence 43%; putative amino acid sequence 35%) (Figure 5). Therefore, the genes required for *chlL* protein activity might activate the *nifH* gene product to obtain Fe protein activity without the requirement for additional genes such as *nifM*, *nifS* or *nifU*. Although regulation of the *chlL*, *N* and *B* genes is not yet clear, using the native *chl* system may provide a strategy for the expression of *nifH* in an active form. In addition, the *nifH* gene product might substitute for the function of *chlL*. The first step was to precisely replace the coding region of *chlL* gene with that of *nifH* gene, keeping the untranslated regulatory regions intact. We designed a strategy to introduce *nifH* into the chloroplast genome by first creating a *petB*::*aadA* insertion mutation and subsequently converting it back to wild-type *petB* with a second homologous recombination event which introduces *nifH*. This event places *nifH* under the control of flanking *chlL*5′ and 3′ regulatory regions (Figure 9). Our results demonstrate that the *nifH* gene product does substitute for the function of *chlL* and it is possible that the chloroplast contains similar ancillary proteins for the biosynthesis of (4Fe-4S) proteins (Cheng et al. 2005a).

One of the next possible approaches would be the replacement of the *nifDK*-like DPOR component *ChlN* and *ChlB* genes with that of Mo-nitrogenase structural *nifDK* genes. Based on the similar assumption that the ChlNB complex may resemble the NifDK complex harboring a similar metal scaffold, provided by yet unknown DPOR biogenesis proteins in chloroplasts, the expectation is to alter the enzymatic DPOR structure towards a functional nitrogenase *in vivo*.

However, it is a major challenge to attempt to engineer a nitrogen-fixing organelle, we still have to face the problem of interfacing plastid physiology with the requirements for nitrogenase activity. As an alternative strategy, one might consider

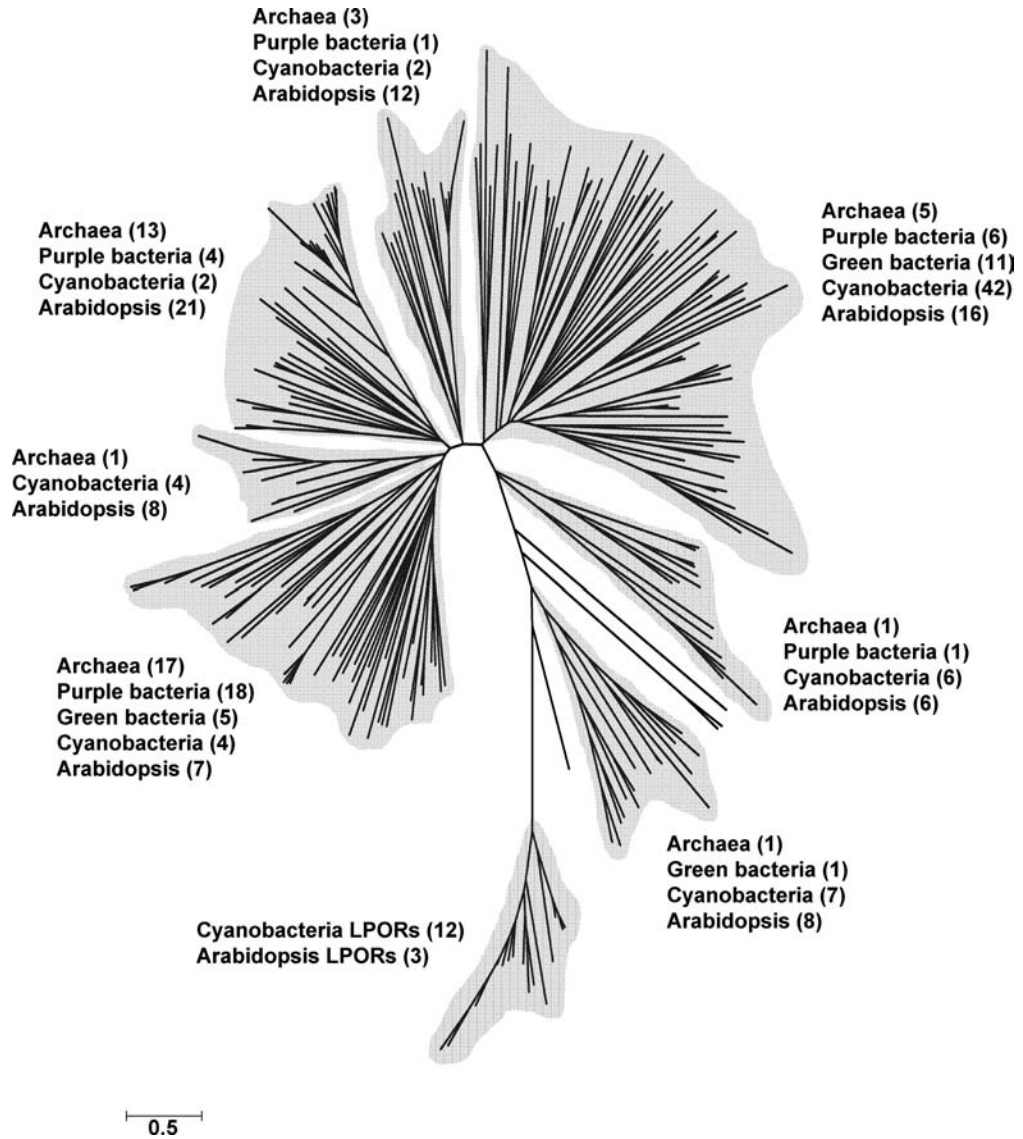

Figure 8. Neighbor joining tree showing the relationships between light-operative protochlorophyllide oxidoreductases (LPORs) and other classical short-chain dehydrogenase/reductases (SDRs).

The species are included in Archaea, purple bacteria, green bacteria and cyanobacteria (Yang and Cheng 2004).

introducing nitrogenase proteins into mitochondria which may provide a suitable energy-rich, reducing environment to support nitrogenase. However, this introduces more complexity because each gene would have to be specifically modified to allow targeting each of the *nif*-encoded proteins into this organelle in the appropriate stoichiometry. Paradoxically, while the technology is available now to introduce and express nitrogenase component proteins in plant cells, substantial progress is limited by gaps in our fundamental knowledge of both plant and microbial physiology.

Emergence of Non-conventional Nitrogenase and Further Predictions

Like everything else on our living planet earth, nitrogenase has been evolving in a dynamic way. Alternative nitrogenases discovered in *Azotobacter vinelandii* using vanadium, iron, instead of molybdenum in an environment lacking molybdenum metal. Both conventional nitrogenase and alternative nitrogenase are two-component complexes and both components are highly sensitive to O_2. *Streptomyces thermoautotrophicus* is recently

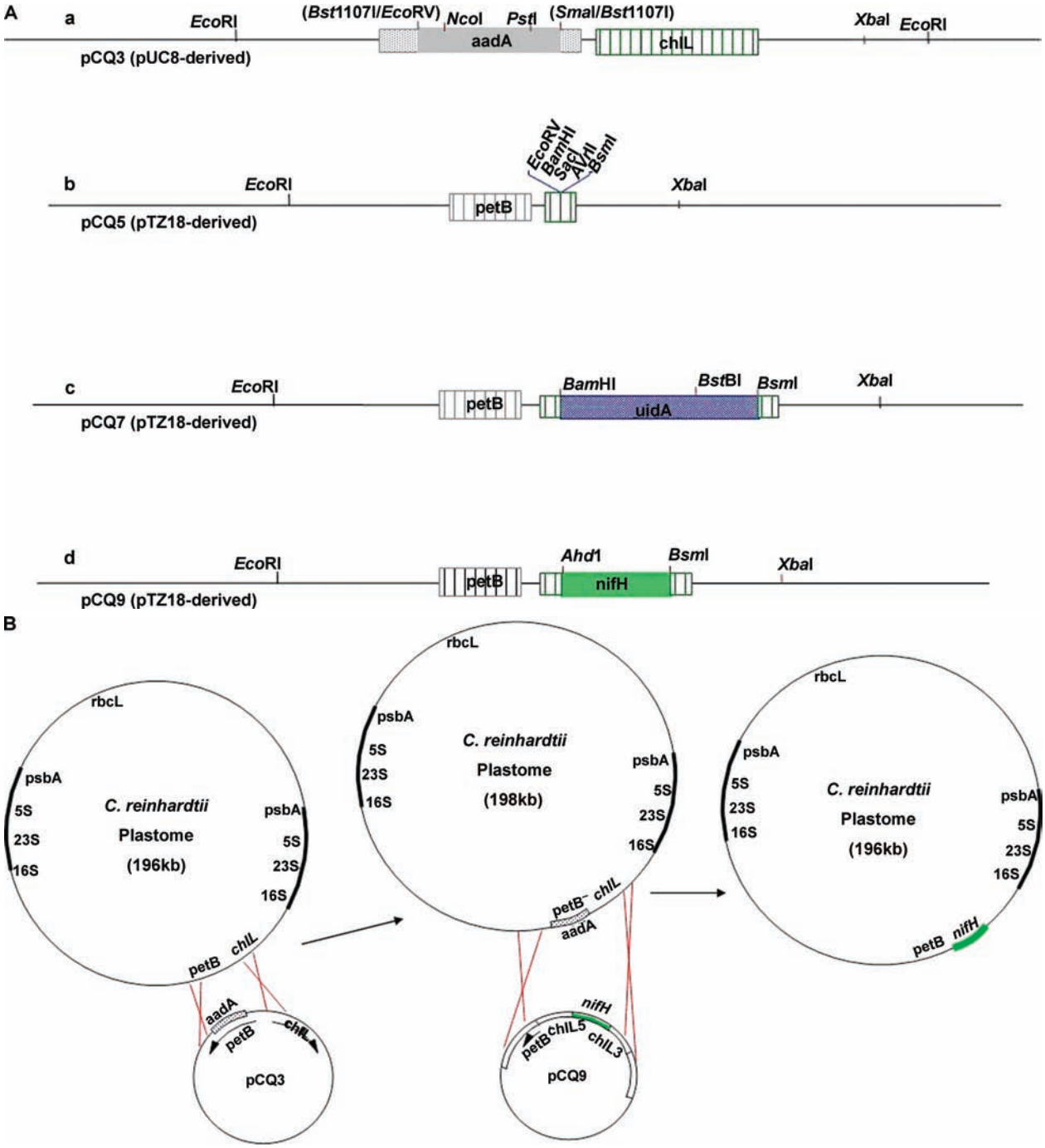

Figure 9. Schematic diagrams of constructs and two-step chloroplast transformation, introduction of nitrogenase Fe protein (NifH) into chloroplast genome of *Chlamydomonas reinhardtii* by replacing its native ChlL.

(A) (a) Chloroplast transformation vector pCQ3. The *aadA* cassette is inserted in the opposite orientation in the *petB* coding region; (b) expression vector pCQ5, containing multiple-cloning-site between *chlL* 5′ and 3′ untranslated regulatory sequences for insertion of foreign target genes; (c) secondary chloroplast transformation vector pCQ7, containing *uidA* gene driven by *chlL* promoter; (d) secondary chloroplast transformation vector pCQ9, containing *nifH* gene driven by *chlL* promoter.

(B) Two-step chloroplast transformation via homologous recombination by bombardment vector pCQ3 to obtain *petB* mutant which was used as a recipient for the secondary transformation by delivering vector pCQ9 bearing *nifH* gene to obtain *C. reinhardtii nifH* transplastomic line. *C. reinhardtii uidA* transplastomic line was also achieved by this strategy (Cheng et al. 2005a).

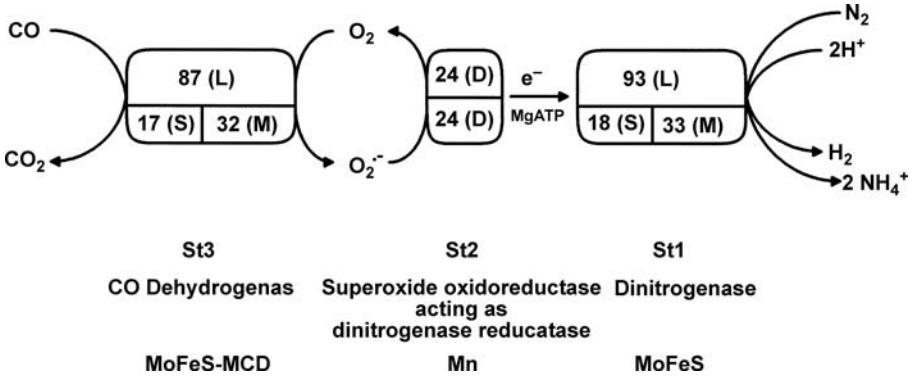

Figure 10. Schematic representation of N_2 fixation in *Streptomyces thermoautotrophicus* (Reproduced from Ribbe et al. (1997) with permission.).

Superoxide is produced by CO dehydrogenase through the oxidation of CO and the transfer of the electrons to O_2. Subsequently, the superoxide is reoxidized by a superoxide oxidoreductase that delivers the electrons to a dinitrogenase. The dinitrogenase is capable of reducing N_2 and H^+, but not ethane (Ribbe et al. 1997).

found to be able to fix dinitrogen, but it harbors a very unusual N_2-fixing system that requires three proteins for nitrogen fixation, a heterotrimeric molybdenum-containing dinitrogenase (St1), a homodimeric manganese-containing superoxide oxidoreductase (St2) and another heterotrimeric molybdenum-containing carbon monoxide dehydrogenase (St3 or CODH) (Figure 10). These proteins differ entirely from the known nitrogenase protein components and show insensitiveness to O_2. Compared to conventional or alternative nitrogenases, the St nitrogenase also requires less ATP: $N_2 + 8H^+ + 8e^- + (4–12)$ MgATP $\rightarrow 2NH_3 + H_2 + (4–12)$ MgADP $+ (4–12)$ Pi.

Previously, *S. thermoautotrophicus UBT1* was isolated from burning charcoal pile. The N_2-fixing ability of *S. thermoautotrophicus* was discovered by growing the strain chemolithoautotrophically with CO or H_2 plus CO_2 under aerobic conditions at 65 °C (Gadkari et al. 1992; Ribbe et al. 1997). That is indeed a good "selection" system where N_2 would be the sole nitrogen source. Another striking characteristic of St nitrogenase is that it is not inhibited by CO, which is the case for the conventional nitrogenase. So far, all known nitrogenases are found in complex prokaryotic versions. There may be more types of prokaryotic nitrogenases with versatile features to be discovered. Because many prokaryotic enzymes do evolve into the eukaryotic version, it would be difficult to rule out the possibility of the existence of a eukaryotic nitrogenase (Yang and Cheng 2004); taking one step further, if eukaryotic nitrogenase does exist in nature, then it may well be utilizing light as an energy source (Cheng et al. 2005b) (LPOR-like nitrogenase: light-utilizing n2ase (LUN)) (Figure 11). The natural history of DPOR and LPOR is a perfect example and indication, driving research to reveal the possible existence of biological nitrogen fixation in wild species of non-legume plants; that is, the possible existence of another non-conventional nitrogenase within the plant kingdom (Figures 8, 11, 12).

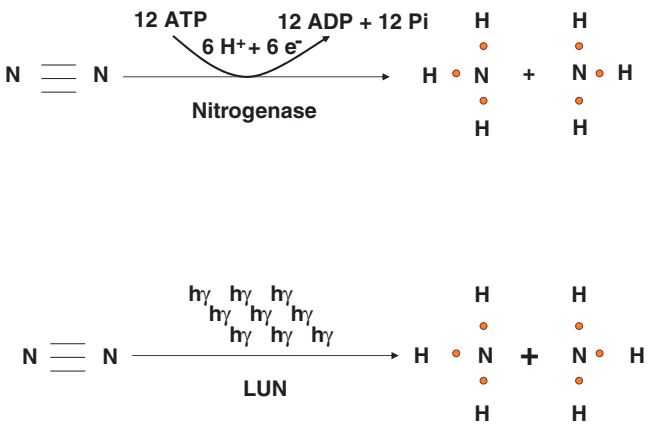

Figure 11. Possibility of the existence of non-conventional nitrogenase.

Conventional nitrogenase complex convert N_2 into NH_3 by using adenosine triphosphate (ATP) as energy; if a light-operative protochlorophyllide oxidoreductase (LPOR)-like enzyme or non-LPOR enzyme could do the same catalysis by a single polypeptide, it could well be a light-utilizing nitrogenase (LUN).

One possibility may be as suggested in Figure 12, showing the relative appearances of nitrogenase, DPOR, LPOR and putative nitrogenase in geologic time (past and future). Such predictions are encouraged by the fact that the anaerobic enzyme has almost certainly been replaced in many reactions by the more efficient and irreversible aerobic version. On the other hand, nature has shown several examples of evolutionary convergence where an enzyme catalyzing a highly specific, O_2-requiring reaction has an oxygen-independent counterpart, able to carry out the same reaction under anoxic conditions (Raymond and Blankenship 2004). While many unpublished experiments have

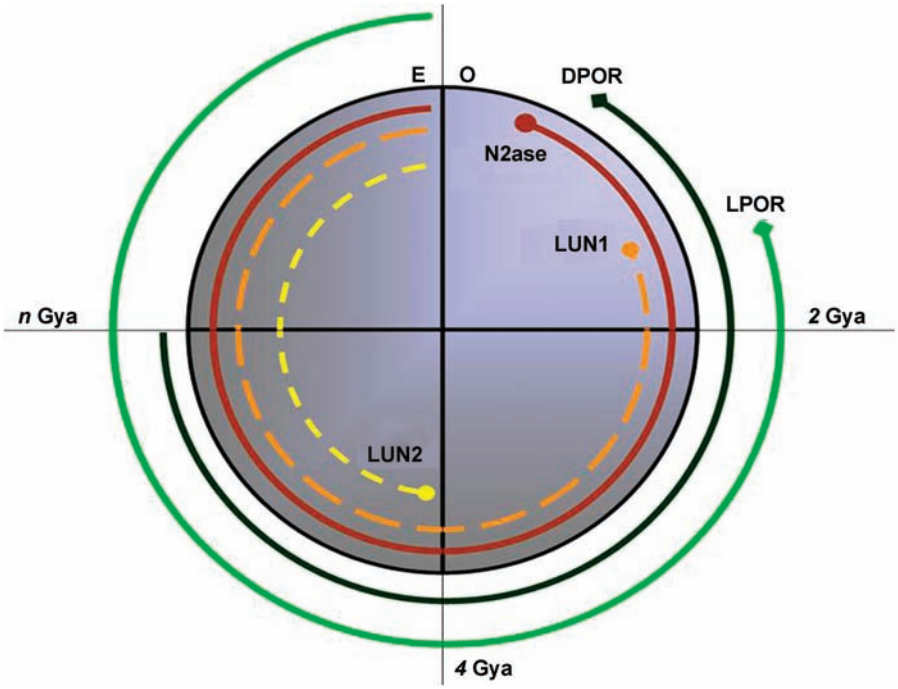

Figure 12. Schematic diagram high-lighting the origin of N2ase, dark-operative protochlorophyllide oxidoreductase (DPOR), light-operative protochlorophyllide oxidoreductase (LPOR) and putative light-utilizing nitrogenase (LUN).

O, origin of our planet earth, formed approximately 4–4.5 Gya ago; E, life of the earth (it has been estimated that it could live on for at least another 4–5 Gya); *n*, any geologic time in the future when some old-fashioned enzymatic systems might be completely wiped out by nature; N2ase appeared at very early stage of an anaerobic world; DPOR was evolved from the same ancestor as n2ase; LPOR appeared when there was considerable O_2 in the atmosphere; LUN1 could have existed around the same period of the origin of LPOR or there may be a possible later evolutionary event giving birth to LUN2 by an unknown driving force in nature or lab.

been conducted to achieve the perfect CO_2/O_2 ratio by engineering Rubisco small subunits (Spreitzer et al. 2007), a better Rubisco enzyme has been found already existing in nature (R. Spreitzer, pers. comm., 2006). Nevertheless, efforts need to be made for an extensive investigation with an open mind and rational searching/selection system, towards finding more enzymes that are able to achieve N_2-fixing for microbes or even a eukaryotic version for plants. Hundreds of genome projects have now been completed, leaving on average more than 50% of discovered proteins waiting to be assigned functions. The speed of genome-wide data accumulation is exponential.

Regarding the possible origin of LUN, might it be the same as LPOR as from SDR protein family, might it be an LPOR, or from a completely different family, or never even have existed at all? Perhaps no plant is able to fix nitrogen? Do plants really not fix nitrogen without a nodule? So far, a definite answer is difficult. On the other hand, if such proteins do exist, a network might already be in place, saving people from having to undergo sophisticated "metabolic engineering"; if not, one may eventually figure out the scientific reason behind it. Either way, it may certainly press science forward in the 21st century.

At this point, one may naively begin investigating the idea of massive overexpression of such proteins heterologously in bacteria. One may never find it active in the sense that current popular methods for assaying nitrogenase are based on C_2H_2 reduction and H_2 evolution. Bearing in mind that, to achieve today's knowledge about microbial N_2-fixing, it took over a hundred years of research in the field of biological nitrogen fixation concentrating mostly on symbiosis models. However, the speed of discovery in modern science should never be underestimated.

Just as B. E. Smith once stated "history tells us that, in this field, the one prediction that we can safely make is that there will be *surprises* awaiting us".

Concluding Remarks

Nitrogen fixation is an "old" topic in scientific terms since it is over a century ago that scientists experimentally proved that some "unique" species of plants, with the help of microbes that were later found bearing nitrogenase, are capable of utilizing

atmospheric nitrogen. The thought of engineering major crops to adapt such a capability, either by association/symbiosis with microbes or by introducing nitrogenase directly into the plant, was proposed several decades ago but only limited experimental approaches were carried out due to the complexity of the nature of the biological nitrogen fixation process. This review summarizes some achievements in revealing the biochemical mechanism of nitrogenase, comparing its evolutionary related proteins, the discovery of non-conventional nitrogenase and initial efforts of introducing nitrogenase into chloroplast. In addition, this review also highlights the new idea of investigating the "likelihood" or "unlikelihood" of the existence of a light-utilizing nitrogenase in the plant kingdom during the "ongoing" course of evolution.

Acknowledgements

I thank many of my colleagues for helpful discussions and apologize for not thanking them individually. This review is dedicated to Professor C. B. You, who sadly passed away in 1994, and will be remembered as one of the early researchers in the field of biological nitrogen fixation in China.

References

Bauer CE, Young DA, Marrs BL (1988). Analysis of the *Rhodobacter capsulatus puf* operon. Location of the oxygen-regulated promoter region and the identification of an additional puf-encoded gene. *J. Biol. Chem.* **263**, 4820–4827.

Bolin TJ, Campobasso N, Muchmore SW, Morgan TV, Mortenson LE (1993). The structure and environment of the metal clusters in the MoFe protein from *Clostridium pasteurianum*. In: Stiefel EI, Coucouvanis D, Newton WE, eds. *Molybdenum Enzymes, Cofactors, and Model Systems*. American Chemical Society, Washington DC. pp. 186–195.

Burke-Aquero DH (1992). *Molecular Genetic and Molecular Evolutionary Studies on the Bacteriochlorophyll Synthesis Genes of Rhodobacter capsulatus*. Ph.D. Thesis. University of California, Berkeley, CA.

Burke DH, Hearst JE, Sidow A (1993). Early evolution of photosynthesis: clues from nitrogenase and chlorophyll iron proteins. *Proc. Natl. Acad. Sci. USA* **90**, 7134–7138.

Brocker MJ, Virus S, Ganskow S, Heathcote P, Heinz DW, Schubert WD et al. (2008). ATP-driven reduction by dark-operative protochlorophyllide oxidoreductase from *Chlorobium tepidum* mechanistically resembles nitrogenase catalysis. JBC M708010200v1 (online) http://www.jbc.org/cgi/content/abstract/M708010200v1

Cheng Q, Day A, Dowson-Day M, Shen GF, Dixon R (2005a). The *Klebsiella pneumoniae* nitrogenase Fe protein gene (*nifH*) functionally substitute for the *chlL* gene in *Chlamydomonas reinhardtii*. *Biochem. Biophys. Res. Commun.* **329**, 966–975.

Cheng Q, Yang J, Day A, Dowson-Day M, Dixon R (2005b). Evolutionary implication of nitrogenase-like genes in plant kingdom and prospects for *nif* gene transfer in model eukaryotes. In: Wang YP et al., eds. *Biological Nitrogen Fixation, Sustainable Agriculture and the Environment*. pp. 387–389.

Dahlin C, Aronsson H, Wilks HM, Lebedev N, Sundqvist C, Timko MP (1999). The role of protein surface charge in catalytic activity and chloroplast membrane association of the pea NADPH:protochlorophyllide oxidoreductase (POR) as revealed by alanine scanning mutagenesis. *Plant Mol. Biol.* **39**, 309–323.

Dixon RA, Postgate JR (1972). Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. *Nature* **237**, 102–103.

Dos Santos PC, Dean DR, Hu Y, Ribbe MW (2004). Formation and insertion of the nitrogenase iron-molybdenum cofactor. *Chem. Rev.* **104**, 1159–1173.

Dowson-Day MJ, Ashurst JL, Watts J, Dixon RA, Merrick MJ (1991). Studies of the potential for expression of nitrogenase Fe-protein in cells of higher plants. In: Polsineli M, Materassi R, Vincenzini M, eds. *Nitrogen Fixation: Proceedings of the 5th International Symposium on Nitrogen Fixation with Non-Legumes*. Kluwer, Dordrecht. pp. 659–669.

Earl CD, Ausubel FM (1982). The genetic engineering of nitrogen fixation. *Technol. Rev.* **85**, 65–71.

Einsle O, Tezcan FA, Andrade SL, Schmid B, Yoshida M, Howard JB et al. (2002). Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. *Science* **297**, 1696–1700.

Fujita Y, Takahashi Y, Shonai F, Ogura Y, Matsubara H (1991). Cloning, nucleotide sequences and differential expression of the *nifH* and *nifH*-like (*frxC*) genes from the filamenous and nitrogen-fixing cyanobacterium *Plectonema boryanum*. *Plant Cell Physiol.* **32**, 1093–1106.

Fujita Y, Takahashi Y, Chungai M, Matsubara H (1992). The *nifH*-like (*frxC*) gene is involved in the biosynthesis of chlorophyll in the filamentous cyanobacterium *Plectonema boryanum*. *Plant Cell Physiol.* **33**, 81–92.

Fujita Y, Matsumoto H, Takahashi Y, Matsubara H (1993). Identification of a *nifDK*-like gene (ORF467) involved in the biosynthesis of chlorophyll in the cyanobacterium *Plectonema boryanum*. *Plant Cell Physiol.* **34**, 305–314.

Fujita Y (1996). Protochlorophyllide reduction: a key step in the greening of plants. *Plant Cell Physiol*. **37**, 411–421.

Fujita Y, Bauer CE (2000). Reconstitution of light-independent protochlorophyllide reductase from purified BchL and BchNBchB subunits. *J. Biol. Chem.* **275**, 23583–23588.

Gadkari D, Morsdorf G, Meyer O (1992). Chemolithoautotrophic assimilation of dinitrogen by *Streptomyces thermoautotrophicus UBT1*: identification of an unusual N_2-fixing system. *J. Bacterial.* **174**, 6840–6843.

Georgiadis MM, Komiya H, Chakrabarti P, Woo D, Kornuc JJ, Rees DC (1992). Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science* **257**, 1653–1659.

Howard JB, Rees DC (2006). How many metals does it take to fix N_2? A mechanistic overview of biological nitrogen fixation. *Proc. Natl. Acad. Sci. USA* **103**, 17088–17093.

Kim J, Rees DC (1992). Crystallographic structure and functional implications of the nitrogenase molybdenum-iron protein from *Azotobacter vinelandii*. *Nature* **360**, 553–560.

Kim J, Woo D, Rees DC (1993). X-ray crystal structure of the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 3.0 Å resolution. *Biochemistry* **32**, 7104–7115.

Lidholm J, Gustafson P (1991). Homologues of the green algal *gidA* gene and the liverwort *frxC* gene are present on the chloroplast genomes of conifers. *Plant Mol. Biol.* **17**, 787–798.

Li JM, Goldschmidt-Clermont M, Timko MP (1993). Chloroplast-encoded *chlB* is required for light-independent protochlorophyllide reductase activity in *Chlamydomonas reinhardtii*. *Plant Cell* **5**, 1817–1829.

Liu XQ, Xu H, Huang C (1993). Chloroplast *chlB* gene is required for light-independent chlorophyll accumulation in *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* **23**, 297–308.

Merrick M, Dixon R (1984). Why don't plants fix nitrogen? *Trends Biotechnol.* **2**, 162–166.

Nomata J, Mizoguchi T, Tamiaki H, Fujita Y (2006). A second nitrogenase-like enzyme for bacteriochlorophyll biosynthesis: reconstitution of chlorophyllide a reductase with purified X-protein (Bchx) and Yz-protein (Bchy-Bchz) from *Rhodobacter capsulatus*. *J. Biol. Chem.* **281**, 15021–15028.

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S et al. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322**, 572–574.

Raymond J, Blankenship R (2004). Biosynthetic pathways, gene replacement and the antiquity of life. *Geobiology* **2**, 199–203.

Rees DC, Kim J, Georgiadis MM, Komiya H, Chirino AJ, Wood D et al. (1993). Crystal-structures of the iron protein and molybdenum iron protein of nitrogenase. *ACS Symposium Series 535*, 170–185.

Ribbe M, Gadkari D, Meyer O (1997). N_2 fixation by *Streptomyces thermotrophicus* involves a molybdenum-dinitrogenase and a manganese-superoxide oxidoreductase that couple N_2 reduction to the oxidation of superoxide produced from O_2 by a molybdenum-CO dehydrogenase. *J. Biol. Chem.* **272**, 26627–26633.

Richards RL, Smith BE (2004). Future challenges and prospects. In: B. E. Smith, R. Richards L and Newton WE, eds. *Catalysts for Nitrogen Fixation: Nitrogenases, Relevant Chemical Models and Commercial Processes*. Kluwer Academic Publishers, Dordrecht. pp. 333–336.

Schindelin H, Kisker C, Schlessman JL, Howard JB, Rees DC (1997). Structure of ADP x AlF4(-)-stabilized nitrogenase complex and its implications for signal transduction. *Nature* **387**, 370–376.

Schoefs B, Franck F (2003). Protochlorophyllide reduction: mechanisms and evolution. *Photochem. Photobiol.* **78**, 543–557.

Spreitzer RJ, Peddi SR, Satagopan S (2007). Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco. *Proc. Natl. Acad. Sci. USA* **102**, 17225–17230.

Suzuki JY, Bauer CE (1995). A prokaryotic origin of light-dependent chlorophyll biosynthesis of plants. *Proc. Natl. Acad. Sci. USA* **92**, 3749–3753.

Yang J, Cheng Q (2004). Origin and evolution of the light-dependent protochlorophyllide oxidase (*LPOR*) genes. *Plant Biol.* **6**, 537–544.

Yang Z, Bauer C (1990). Rhodobacter capsulatus genes involved in the early steps of the bacteriochlorophyll pathway. *J. Bacteriol.* **172**, 5001–5010.

Zsebo KM, Hearst JE (1984). Genetic physical mapping of a photosynthetic gene cluster from *R. capsulatus*. *Cell* **37**, 937–947.

(Handling editor: Chun-Ming Liu)