J Mol Evol (1989) 29:422–428

Journal of Molecular Evolution

Glutamine Synthetase II in *Rhizobium*: Reexamination of the Proposed Horizontal Transfer of DNA from Eukaryotes to Prokaryotes

Robert G. Shatters[1]* and Michael L. Kahn[2,3]

[1]Program in Genetics and Cell Biology, [2]Institute of Biological Chemistry, and [3]Department of Microbiology, Washington State University, Pullman, Washington 99164-6340, USA

Summary. We have determined the DNA sequence of a *Rhizobium meliloti* gene that encodes glutamine synthetase II (GSII). The deduced amino acid sequence was compared to that of *Bradyrhizobium japonicum* GSII and those of various plant and mammalian glutamine synthetases (GS) in order to evaluate a proposal that the gene for this enzyme was recently transferred from plants to their symbiotic bacteria. There is 83.6% identity between the *R. meliloti* and *B. japonicum* proteins. The bacterial GSII proteins average 42.5% identity with the plant GS proteins and 41.8% identity with their mammalian counterparts. The plant proteins average 53.7% identity with the mammalian proteins. Thus, the GS proteins are highly conserved and the divergence of these proteins is proportional to the phylogenetic divergence of the organisms from which the sequences were determined. No transfer of genes across large taxonomic gaps is needed to explain the presence of GSII in these bacteria.

Key words: Glutamine synthetase — *Rhizobium* — *Bradyrhizobium* — Protein evolution

Introduction

Glutamine is an important precursor in many biosynthetic pathways. It is made by glutamine synthetase (GS), which combines glutamate and ammonium in an ATP-dependent reaction. In addition, GS is the first enzyme in the glutamine synthetase–glutamate synthase pathway used by both prokaryotes and eukaryotes to assimilate ammonium. Because the synthesis and activity of prokaryotic GS are affected by complex regulatory controls (for review see Gussin et al. 1986), GS is likely to be a key enzyme in the control of carbon and nitrogen metabolism.

Two types of GS enzymes have been studied in some detail. One is a dodecamer of identical subunits that has been found only in prokaryotes (Stadtman and Ginsburg 1974; Darrow 1980; Colombo and Villafranca 1986). The second is an octamer of identical subunits and has been found in animals (Hayward et al. 1986; Gibbs et al. 1987), higher plants (Gebhardt et al. 1986; Tischer et al. 1986; Tingey and Coruzzi 1987; Tingey et al. 1987), and a small group of bacteria, including most species of *Rhizobium* (Darrow and Knotts 1977) and *Agrobacterium* (Fuchs and Keister 1980). This group of bacteria also contains the dodecameric form of GS. To date, all bacteria known to contain the gene for the octameric GS, *glnII*, also are involved in symbiotic relationships with higher plants. An obvious question is whether the existence of the eukaryotic type of GS enzyme in these bacteria is related to their association with eukaryotic hosts.

Carlson and Chelm (1986) recently determined the DNA sequence of the *glnII* gene from *Bradyrhizobium japonicum*, and showed a substantial similarity between the bacterial sequence and a number of plant GS sequences. Because of the limited distribution of the enzyme in prokaryotes, they

* *Current address:* USDA Northern Regional Research Center, Peoria, Illinois 61604, USA
Offprint requests to: M.L. Kahn

suggested that *glnII* occurs in plant-associated bacteria as the result of a transfer of the gene from the plant to the bacteria. As *B. japonicum* is a member of a group of bacteria that invade the root tissue of leguminous plants and multiply in specialized plant root cells, it seems plausible that the endosymbiotic relationship is intimate enough to allow the exchange of genetic material. Confirmation of this speculation would be significant, because it would be direct evidence that genetic exchanges can occur from eukaryotes to prokaryotes.

This paper presents the *glnII* sequence from *Rhizobium meliloti,* another bacterium symbiotic with legumes (*Rhizobium* and *Bradyrhizobium,* formerly considered to be congeneric, recently were separated taxonomically). We also present a more comprehensive comparison of bacterial, plant, and animal GS sequences. Our conclusions do not support the idea that *glnII* is of plant origin.

Materials and Methods

Bacterial Strains and Plasmids. Rhizobium meliloti 104A14 has been described by Somerville and Kahn (1983). *Escherichia coli* strain JM105 (Yanisch-Perron et al. 1985) was used for recombinant work. Plasmid pJS73 (Somerville et al. 1989) contained the entire *glnII* gene and was used as the source of *glnII* DNA.

Subcloning and Sequencing. After mapping the appropriate restriction sites in the *glnII* locus, we cloned various restriction fragments into the M13 vectors. The DNA sequence of these fragments then was determined using modified bacteriophage T7 DNA polymerase (Sequenase) and the dideoxynucleotide chain termination kit supplied by U.S. Biochemicals. Both strands of the *glnII* coding region were sequenced.

The *glnII* subsequences were compiled on a Kaypro IV personal computer using the Pustell sequence analysis software (Pustell and Kafatos 1982). Subsequent sequence comparisons, alignments, and data base searches were done with the University of Wisconsin sequence analysis software (Devereux et al. 1984) run on a VAX 11/785 and maintained by the Washington State University VADMS Center.

Results and Discussion

Comparison of the R. meliloti glnII Sequence with Those of Other Organisms

The *glnII* gene was cloned from *R. meliloti* 104A14 *glnA* mutant, GLN210, by isolating a glutamine auxotroph using nitrous acid mutagenesis and subsequently rescuing the mutant with an *R. meliloti* cosmid gene bank (Somerville et al. 1989). Cosmid pJS73 was shown to contain the *glnII* structural gene. The DNA sequence of this gene was determined as described in Materials and Methods (Fig. 1). Also shown in Fig. 1 are the derived protein sequence and the nucleotides of the *B. japonicum glnII* sequence that differ from the *R. meliloti glnII* gene.

In order to test the speculation of Carlson and Chelm (1986) that the bacterial *glnII* gene arose by transfer of DNA from the host plant to the bacteria, we did a more inclusive comparison using the deduced amino acid sequences of GS from symbiotic bacteria and all the available GS sequences from various plants and animals. Table 1 shows the results of pairwise comparisons of the sequences in which the percentage of identical amino acids (upper sector) or functionally similar amino acids (lower sector) was determined. The percent identity between the bacterial GSII sequences and the eukaryotic sequences averages 42.5%, and the percent similarity averages 60.6%. These values represent significant homology between the bacterial glutamine synthetase II (GSII) and the eukaryotic GS sequences (Doolittle 1986). Interestingly, the values derived from comparing the bacterial GSII sequences to the plant sequences are virtually the same as those derived from comparing the bacterial to the mammalian sequences. Because the homology between the plant and animal GS sequences averages 53.7% identity, the similarity between the bacterial–plant and the bacterial–mammalian GS comparisons suggests that the bacterial sequence diverged from the eukaryotic sequence prior to the separation of plants from animals.

Similar Homology of the Bacterial GSII Proteins to Both Plant and Animal GS Sequences Is Not Due to Uneven Rates of Evolution

It is conceivable that a faster rate of change of the *glnII* gene in the bacteria, after lateral transfer, might explain some of these data. If the bacterial *glnII* gene arose as a descendant of the plant gene, then we would expect that in those amino acids where the bacterium matches only the plant or only the mammalian sequence, it should most often match that of the plant. The GS sequences were aligned as described in Materials and Methods using consensus sequences derived from bacteria, plants, and mammals (Fig. 2). Among all the taxonomic groups, similarities of the GS proteins were observed throughout the sequences, with definite clustering of conserved amino acids, probably at functionally significant regions. The comparison also shows 20 positions where the bacterial sequence matches the plant sequence but not the mammalian one, and 21 where the bacterial sequence matches the mam-

```
        MetThrLysTyrLysLeuGluTyrIleTrpLeuAspAlaThrArgProTyrGlnThrLeuArgGlyLysThrGlnIleLysGluPheAspAlaPheProThrLeuGluGlnLeuProLeu
Rm-glnII ATGACCAAGTATAAGCTCGAGTACATCTGGCTTGATGCTACACGCCCGTACCAAACTCTCCGCGGAAAAACGCAGATCAAGGAATTCGACGCGTTTCCGACGCTCGAGCAGCTTCCGCTC  120
Bj-glnII                                C  C GATATACG   ACT CG ACT G    C    T              CGT   C             CG

        TrpGlyPheAspGlySerSerThrLeuGlnAlaGluGlyArgThrSerAspCysValLeuLysProValThrValTyrProAspProValArgThrAsnGlyAlaLeuValMetCysGlu
Rm-glnII TGGGGCTTTGACGGCAGCTCTACGCTGCAGGCCGAGGGGCGCACGTCCGATTGTGTGCTGAAGCCGGTTACCGTCTATCCCGATCCGGTCCGCACGAATGGCGCATTGGTCATGTGCGAG  240
Bj-glnII        C T  TC  C C A        A C A  GC  T    C          A CG     TC  G  CG C CG    C C   TGC C  G       A

        ValMetMetProAspAlaGluThrProHisAlaSerAsnThrArgAlaThrValLeuAspAspGluGlyAlaTrpPheGlyPheGluGlnGluTyrPhePheTyrLysAsnGlyArgPro
Rm-glnII GTGATGATGCCGGACGCAGAGACACCGCATGCGTCGAATACCCGGGCAACGGTCCTCGATGACGAAGGCGCCTGGTTCGGCTTCGAGCAGGAATATTTCTTCTACAAGAACGGCCGCCCA  360
Bj-glnII  C       C  T GCA   C    C  C  C  C AG  C  C  CA T    C   CC                              C          G      T  G

        LeuGlyPheProGluGlnGlyTyrProAlaProGlnGlyProTyrTyrThrGlyValGlyTyrLysAsnValGlyAspValAlaArgGlnIleValGluGluHisLeuAspIleCysLeu
Rm-glnII CTGGGCTTCCCGGAGCAGGGCTATCCGGCGCCGCAGGGCCCATACTATACCGGCGTCGGCTACAAGAATGTCGGCGACGTCGCGCGCCAGATCGTTGAAGAGCATCTCGACATCTGCCTT  480
Bj-glnII  C         ACCTCC  T                   G    C           T TC  C        G  C   A      C             C      C

        AlaAlaGlyIleAsnHisGluGlyIleAsnAlaGluValAlaLysGlyGlnTrpGluPheGlnIlePheGlyLysGlySerLysLysAlaAlaAspGluValCysValAlaArgTyrLeu
Rm-glnII GCTGCGGGCATCAACCACGAAGGCATCAACGCCGAAGTGGCCAAGGGGCAATGGGAGTTCCAGATCTTTGGCAAGGGTTCCAAAAAAGCTGCCGACGAGGTCTGCGTGGCGCGCTATCTC  600
Bj-glnII  G  C          T             G    C  G    C  G    A          C       C    G  G  C  T    AA G  GA    C    C  G

        LeuValArgLeuThrGluLysTyrGlyIleAspValGluPheHisCysLysProLeuGlyAspThrAspTrpAsnGlySerGlyMetHisAlaAsnPheSerThrAlaTyrLeuArgGlu
Rm-glnII CTCGTGCGCCTGACGGAAAAATACGGCATCGACGTCGAATTCCACTGCAAGCCGCTCGGCGACACGGACTGGAACGGCTCGGGCATGCACGCGAACTTCTCGACGGCCTATCTGCGTGAA  720
Bj-glnII A GC          C  G  G          A    G               T                    C          C          C AG  CA    CACG

        ValGlyGlyGlnAspTyrPheGluAlaLeuMetAlaAlaPheGluLysAsnLeuHisAspHisIleAsnValTyrGlyProAspAsnHisLeuArgLeuThrGlyLysHisGluThrAla
Rm-glnII GTCGGCGGCCAGGACTATTTCGAAGCGCTGATGGCAGCCTTCGAGAAGAACCTGCACGACCACATCAACGTCTATGGCCCTGATAACCATTTGCGCCTGACTGGCAAGCACGAGACGGCG  840
Bj-glnII         A   G C    G          T       C        ATG        GC     C    G C   G CAA  T    C             C

        ProTrpAspLysPheSerTyrGlyValAlaAspArgGlyAlaSerIleArgValProHisSerPheValAsnAsnAlaTyrProGlyTyrLeuGluAspArgArgAlaAsnSerGlnGly
Rm-glnII CCGTGGGATAAGTTCAGCTACGGCGTGGCAGATCGCGGCGCCTCTATCCGCGTGCCCCATAGTTTCGTCAACAATGCGTATCCGGGCTATCTCGAAGATCGCCGCGCGAACTCCCAGGGC  960
Bj-glnII       A C                   C  C       T  G          G  CTCC          C GC  CAA        G     C  T   C       G

        AspProTyrGlnMetLeuLeuSerSerEnd
Rm-glnII GACCCCTACCAGATGCTTCTATCATCCTGA  990
Bj-glnII      A        CG   GCAG
```

Fig. 1. Sequence of the *glnII* gene for *Rhizobium meliloti* 104A14. The top line represents the three-letter amino acid code for the protein and the middle line is the entire *glnII* gene nucleic acid sequence starting with the ATG start codon. The bottom line shows only the nucleotide bases of the *Bradyrhizobium japonicum* sequence that differ from the *R. meliloti* sequence.

Table 1. Amino acid comparison of the eukaryotic and prokaryotic glutamine synthetases

	Pairwise comparison									
Amino acid sequences	*H.s.*	*C.l.*	*M.s.*	*P.v.*	*P.v.* GSr1	*P.s.* GSr2	*N.p.*	*R.m.*	*B.j.*	*E.c.*
Homo sapiens GS[a]	—	93.6	54.1	54.5	53.4	53.7	52.5	41.5	42.2	16.7
Cricetulus longicaudatus GS[b]	97.0	—	54.1	54.5	53.9	54.3	52.5	40.6	42.9	15.6
Medicago sativa GS[c]	70.1	69.7	—	92.4	88.8	88.4	91.6	42.4	43.8	13.7
Phaseolus vulgaris GSr1[d]	70.0	69.8	96.9	—	87.7	87.3	87.6	40.0	42.2	18.7
Phaseolus vulgaris GSr2[e]	70.4	69.2	94.4	94.1	—	99.7	88.2	42.4	43.8	11.7
Pisum sativum GS[f]	70.4	69.6	94.1	94.1	99.7	—	88.2	42.4	43.8	11.8
Nicotiana plumbaginifolia GS[g]	68.4	68.9	95.8	94.1	94.1	94.1	—	42.1	42.9	19.6
Rhizobium meliloti GSII[h]	61.9	61.9	60.4	60.4	58.8	58.9	59.8	—	83.6	17.6
Bradyrhizobium japonicum GSII[i]	60.6	60.6	61.5	61.8	60.9	60.6	60.9	90.9	—	19.0
Escherichia coli GS[j]	44.5	42.1	42.2	41.5	37.8	42.8	47.3	41.4	40.3	—

The top right sector shows the % identity. The bottom left sector shows the % similarity based on categorizing the amino acids into functionally similar groups

[a] Gibbs et al. (1987). [b] Hayward et al. (1986). [c] Tischer et al. (1986). [d,e] Gebhardt et al. (1986). [f] Tingey et al. (1987). [g] Tingey and Coruzzi (1987). [h] This paper. [i] Carlson and Chelm (1986). [j] Colombo and Villafranca (1986)

```
                                      P     M                 M
Mammalian Consensus  M-TSASSHLN K-IKQ-Y--L PQGEKVQAMY IW-DGTGEGL RCKTRTLD-E
    Plant Consensus  .......MSL LSDLINLNLS ESTEKIIAEY IWIGGSGMDL RSKART....
Bacterial Consensus  .......... .......... ..MTKYKLEY IWLD---P-- -LRGKTQIKE
Overall Consensus                             K    Y IW              T

                     51                                                  100
                         ?       M       M         PM     M               ? P
Mammalian Consensus  ......PKCV EELPEWNFDG SST-QSEGSN SDMYL-P-AM FRDPFR-DPN
Plant Consensus      .LPGPV-DP. AKLPKWNYDG SSTGQAPG-D SEVILYPQAI FKDPFRRGNN
Bacterial Consensus  F--FP...TL E--PLWGFDG SST-QAEG-- SDCVLKPV-V -PD---RTNG
Overall Consensus                  P W  DG SST Q  G   S   L P      D

                     101                                                 150
                          MM  P               ??                M
Mammalian Consensus  KLV-CEV..F KYNR-PAETN LRH-CKRIMD MVSNQHPWFG MEQEY.TLMG
Plant Consensus      ILVICDAYTP AGEPIPTNKR H-AAKIFSHP DVVAEVPWYG IEQEYTLLQK
Bacterial Consensus  -LVMCEVMMP D--......-- TPHASN-RAT -LDD-GAWFG FEQEY..FFY
Overall Consensus     LV C                                  W G  EQEY

                     151                                                 200
                       M  P                                    M       M   PP
Mammalian Consensus  TDGHPFGWPS NGFPGPQGPY YCGVG..AD- AYGRDIVEAH YRACLYAGVK
Plant Consensus      DVNWPLGWP- GGFPGPQGPY YCGVG..ADK AFGRDIVDAH YKACVYAGIN
Bacterial Consensus  K-GRPLGFP- -GYPAPQGPY YTGVG--NVG DVAR-IVEEH LD-CLAAGIN
Overall Consensus        PLG P   G P PQGPY Y GVG         R IV  H    C  AG

                     201                                                 250
                        P M     P     M        P  P      P     P
Mammalian Consensus  I-GTNAEVMP AQWEFQIGPC EGI-MGDHLW VARFILHRVC EDFGVIATFD
Plant Consensus      ISGINGEVMP GQWEFQVGPS VGISAGDEVW VARYILERIT E-AGVVVSFD
Bacterial Consensus  HEGINAEVAK GQWEFQIFGK GSKKAADE-- -ARYL--RLT EKYGID-EFH
Overall Consensus      G N EV    QWEFQ           D     AR    R   E  G    F

                     251                                                 300
                           P          M        P      M        M
Mammalian Consensus  PKP.IPGNWN GAGCHTNFST KAMREENGLK -IEEAIEKLS KRH-YHIRAY
Plant Consensus      PKP.IPGDWN GAGAHTNYST KSMREDGGYE VIKKAIEKLG LRHKEHIAAY
Bacterial Consensus  CKPLGDTDWN GSGMHANFST -Y-R-VGG-- YFEALMAAF- KNL-DHI-VY
Overall Consensus     KP     WN G G H N ST    R   G                  HI  Y

                     301                                                 350
                          M    ?       P      ?       P  M PM    M
Mammalian Consensus  DPKGGLDNAR -LTGFHETSN INDFS-GVAN RSASIRIP.R TVGQEKKGYF
Plant Consensus      ....GEGNER RLTGRHETAD INTF-WGVAN RGASVRV.GR DTEKAGKGYF
Bacterial Consensus  ....GPDN-- RLTGKHETAP W-KFSYGVAD RGASIRVPHS FVNN-Y-GYL
Overall Consensus        G  N    LTG HET      F  GVA  R AS R           GY

                     351                                386
                          P     P
Mammalian Consensus  EDRRPSANCD PF-VTEA--R TCLLNETGDE PFQYKN
Plant Consensus      EDRRPASNMD PYVVTSMIAD TTILWKP... ......
Bacterial Consensus  EDRR-NSQGD PYQ----S-. .......... ......
Overall Consensus    EDRR     D P
```

Fig. 2. Alignment of the glutamine synthetase amino acid consensus sequences from mammals, plants, and bacteria. The three separate consensus sequences were derived from comparisons of the sequences within each group using the GAP program from the Wisconsin software package described in Materials and Methods. For the plant consensus an amino acid code was assigned when three out of five sequences matched. The overall consensus shows those residues that are identical in all three of the above consensuses. A "-" represents a position where the comparison of the residues did not meet the criteria to produce a consensus amino acid. A "." represents a gap placed in the consensus to optimize the sequence alignment. A "P" above the sequences indicates the position of a match between only the bacterial and the plant consensuses. An "M" above the sequences indicates the position of a match between only the bacterial and the mammalian consensuses. A "?" above the sequences indicates a position where the bacterial sequence matches only one of the eukaryotic sequences, but was not used in the analysis because it may be either a result of the gapping program or due to the lack of a consensus in one of the eukaryotic sequences.

malian sequence but not the plant. Again this result is inconsistent with the lateral DNA transfer hypothesis.

Evidence against a faster rate of gene evolution in a prokaryotic background has been presented by Ochman and Wilson (1987). Their data suggest that although silent substitutions in bacterial sequences occur at rates similar to those seen in eukaryotes, the rate of substitutions at replacement sites is significantly slower in prokaryotes than in eukaryotes.

Sequence Divergence of the Eukaryotic GS and the Prokaryotic GSII Is Congruent with Organismal Relationships

Using the data shown in Table 1, we constructed a dendrogram to describe the relationship among the various GS proteins (Fig. 3A). This dendrogram is congruent with the phylogenetic tree of the organisms (Fig. 3B). It is therefore apparent from this work that the homology between the GS genes of eukaryotes and prokaryotes is not due to a recent lateral transfer of DNA, but more likely is due to selective pressure favoring the conservation of a sequence originally present in these bacteria and a primitive eukaryote.

The glnII Genes are More Similar between R. meliloti and B. japonicum than Are the nifH Genes

In light of the high level of GS sequence conservation between eukaryotes and prokaryotes, we looked at the degree of similarity between the two bacterial sequences. The two genes share 90.2% nucleic acid similarity, and the deduced amino acid sequences share 83.6% identity and 90.9% similarity (Table 1). These values were compared to those for the other known amino acid sequences from *R. meliloti* and two *Bradyrhizobium* species [either *B. japonicum* or *B.* sp. (Parasponia)]: *nifH, nodD, nodA,* and *nodB* (Fig. 4). The *glnII* sequences are the most similar, even more similar than the *nifH* genes (Ruvkun and Ausubel 1980).

This similarity between the two prokaryotic *glnII* sequences could be the result of either lateral transfer between bacterial species or evolutionary conservation. Hennecke et al. (1985) showed for three nitrogenase genes (*nifH, nifD,* and *nifK*) that the divergence between bacterial species was proportional to the divergence of 16S RNA. Their result did not support the idea of a recent lateral transfer of the nitrogenase genes. Because the dendrograms for *glnII* match the taxonomic relationships of the organisms from which the GS genes were cloned (Fig. 3), and the across-species similarity for *glnII* is similar to that of *nifH,* the same argument may hold for GSII.

For the dodecameric GS, the active site of the protein is formed at the interface of the subunits (Almassy et al. 1986). If a similar subunit interaction exists with the octameric form, there may be little tolerance for alterations in the subunit interactions and hence in the primary sequence. The strong similarity between animal and plant GS sequences is consistent with this interpretation.

The idea that GSII is a highly conserved protein suggests a specific requirement for its enzymatic activity that cannot be carried out as efficiently by other enzymes. However, *R. meliloti* and *B. japonicum* strains that lack GSII are able to form effective nodules (Somerville et al. 1989; Carlson et al. 1987, respectively), and *Agrobacterium* strains that lack GSII apparently are fully virulent (Rossbach et al. 1988). We have not been able to find any growth deficiency caused by a *glnII* mutation in *R. meliloti* 104A14 using various carbon and nitrogen sources. Because wild type bacteria contain a second functional GS enzyme, it is not clear what role GSII plays in their life history that would account for its evolutionary stability.

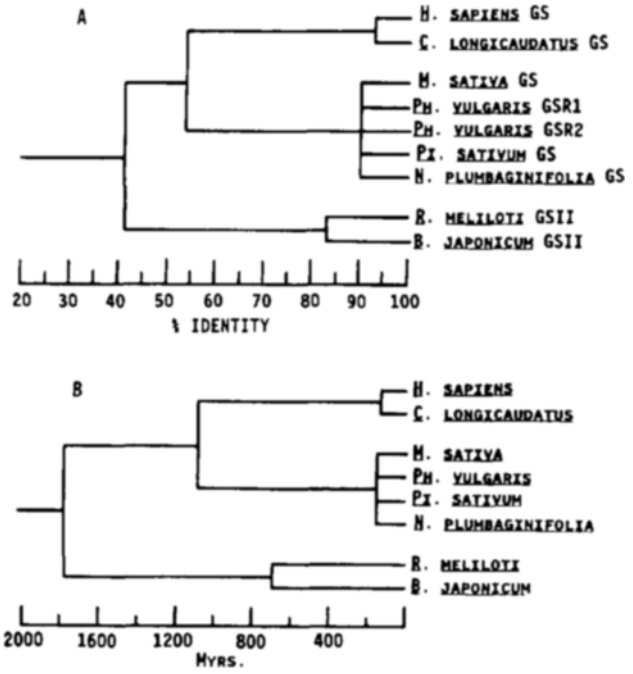

Fig. 3. **A** Dendrogram of amino acid similarities, comparing the eukaryotic GS sequences and the prokaryotic GSII sequences. **B** Dendrogram of the phylogenetic relationships among the species for which the GS sequences were used in Fig. 3A. Evolutionary time points were taken from McKenna (1975), for the divergence of humans from hamsters; Crepet and Taylor (1985) and Lim and Burton (1983), for the appearance of legumes (the evolutionary distance between the plant sequences is not shown); Hasegawa et al. (1985), for the divergence of plants and animals; Kimura and Ohta (1973), for the divergence of eukaryotes and prokaryotes; and Ochman and Wilson (1987), for the divergence of *Bradyrhizobium japonicum* and *Rhizobium meliloti.*

Possible Explanations for the Limited Distribution of the glnII Gene in the Prokaryotes

Having argued that the data do not support the notion that *glnII* exists in plant-associated bacteria because of a recent lateral transfer, we must confront the question of why only these bacteria have this gene. The first possibility is that other bacteria not included in the original surveys actually have the gene. Recently, for example, the gram-positive ni-

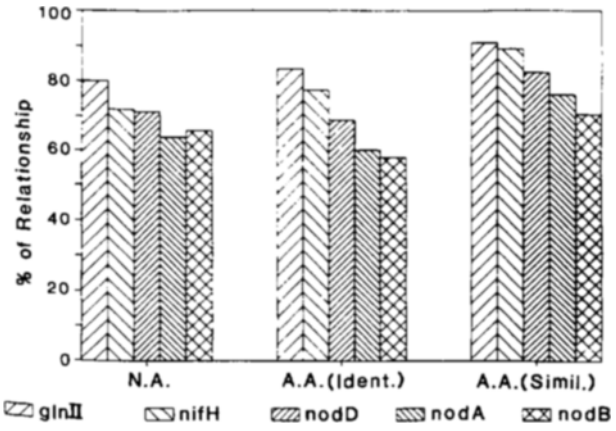

Fig. 4. Extent of sequence similarity between homologous genes from *Rhizobium meliloti* and one of two *Bradyrhizobium* species. N.A. represents nucleic acid similarity and A.A. represents amino acid similarity either for exact matching ("identity") or by categorizing functionally similar amino acids ("similarity"). The comparisons are ▨ GS genes; ▧ *nifH* genes; ▨ *nodD* genes; ▧ *nodA* genes; and ▩ *nodB* genes. The sequences were obtained from *B. japonicum* GSII, Carlson and Chelm (1986); *B. japonicum nifH*, Fuhrmann and Hennecke (1984); *B.* sp. (Parasponia) *nodA, nodB, and nodD*, Scott (1986); *R. meliloti nifH*, Torok and Kondorosi (1981); and *R. meliloti nodA, nodB*, and *nodD*, Egelhoff et al. (1985).

trogen-fixing symbiont *Frankia* has been shown to contain a second glutamine synthetase (Edmands et al. 1987) and DNA that hybridizes to *B. japonicum glnII* DNA (Benson et al. 1988). It is possible that the *glnII* gene was once widely distributed, but that it has been replaced during prokaryotic evolution by *glnA*, perhaps because of the variety of control mechanisms that exist to regulate the enzyme produced by *glnA*. Under this explanation, we would expect to find *glnII* genes in other bacteria. The strong conservation between the *glnII* genes described here suggests that a survey of various bacteria looking for hybridization to *glnII* DNA would be worthwhile in order to define the distribution of this gene better.

A second possibility is that the *glnII* gene, of prokaryotic origin, was transferred to eukaryotes from an endosymbiotic purple bacterium, a group to which *Rhizobium* and *Bradyrhizobium* belong. It has been argued that mitochondria are derived from this group of bacteria (Yang et al. 1985). If *glnII* was present in the protomitochondrion, then transfer of the information for *glnII* synthesis to the nucleus might have occurred during the development of eukaryotes. In the dendrogram of Fig. 3B, this protomitochondrion would be associated with a protoeukaryote that existed before the separation of plants and animals. Therefore, the point of evolutionary divergence of the prokaryotic *glnII* and the eukaryotic GS sequences would be near the point of divergence of the Rhizobiaceae and the mitochondrial ancestor.

Acknowledgments. We thank Linda Moore for critical reading of the manuscript. This research was supported by US Department of Agriculture Competitive Research Grants Office.

References

Almassy RJ, Janson CA, Hamlin R, Xuong N-H, Eisenberg D (1986) Novel subunit–subunit interactions in the structure of glutamine synthetase. Nature (London) 323:304–309

Benson DR, Noridge NA, Tsai YL, Taylor DA (1988) Enzymes of ammonia assimilation in hyphae and vesicles of *Frankia* sp. strain CpI1. In: Bothe H, de Bruijn FJ, Newton WE (eds) Nitrogen fixation: hundred years after. Proceedings of the 7th International Congress on Nitrogen Fixation. Gustav Fischer, New York, p 693

Carlson TA, Chelm BK (1986) Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. Nature (London) 322:568–570

Carlson TA, Martin GB, Chelm BK (1987) Differential transcription of the two glutamine synthetase genes of *Bradyrhizobium japonicum*. J Bacteriol 169:5861–5866

Colombo G, Villafranca JJ (1986) Amino acid sequence of *Escherichia coli* glutamine synthetase deduced from the DNA nucleotide sequence. J Biol Chem 261:10587–10591

Crepet WL, Taylor DW (1985) The diversification of the Leguminosae: first fossil evidence of the Mimosoideae and Papilionoideae. Science 228:1087–1089

Darrow RA (1980) Role of glutamine synthetase in nitrogen fixation. In: Mora J, Palacios R (eds) Glutamine metabolism, enzymology, and regulation. Academic Press, New York, pp 139–166

Darrow RA, Knotts RR (1977) Two forms of glutamine synthetase in free-living root-nodule bacteria. Biochem Biophys Res Commun 78:554–559

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12(1):387–395

Doolittle RF (1986) Of urfs and orfs: a primer on how to analyze derived amino acid sequences. University Science Books, Mill Valley CA

Edmands J, Noridge NA, Benson DR (1987) The actinorhizal root-nodule symbiont *Frankia* sp. strain CpI1 has two glutamine synthetases. Proc Natl Acad Sci USA 84:6126–6130

Egelhoff TT, Fisher RF, Jacobs TW, Mulligan JT, Long SR (1985) Nucleotide sequence of *Rhizobium meliloti* 1021 nodulation genes: *nodD* is read divergently from *nodABC*. DNA 4:241–248

Fuchs RL, Keister DL (1980) Identification of two glutamine synthetases in *Agrobacterium*. J Bacteriol 141:996–998

Fuhrmann M, Hennecke H (1984) *Rhizobium japonicum* nitrogenase Fe protein gene (*nifH*). J Bacteriol 158:1005–1011

Gebhardt C, Oliver JE, Forde BG, Saarelainen R, Miflin BJ (1986) Primary structure and differential expression of glutamine synthetase genes in nodules, roots and leaves of *Phaseolus vulgaris*. EMBO J 5:1429–1435

Gibbs CS, Campbell KE, Wilson RH (1987) Sequence of a human glutamine synthetase gene. Nucleic Acids Res 15:6293

Gussin GN, Ronson CW, Ausubel FM (1986) Regulation of nitrogen fixation genes. Annu Rev Genet 20:567–591

Hasegawa M, Iida Y, Yano T-A, Takaiwa F, Iwabuchi M (1985) Phylogenetic relationships among eukaryotic kingdoms inferred from ribosomal RNA sequences. J Mol Evol 22:32–38

Hayward BE, Hussain A, Wilson RH, Lyons A, Woodcock V, McIntosh B, Harris TJR (1986) The cloning and nucleotide sequence of cDNA for an amplified glutamine synthetase gene from the Chinese hamster. Nucleic Acids Res 14:999–1007

Hennecke HK, Kaluza B, Stackebrandt E (1985) Concurrent evolution of nitrogenase genes and 16S rRNA in *Rhizobium* species and other nitrogen fixing bacteria. Arch Microbiol 142:342–348

Kimura M, Ohta T (1973) Eukaryotes–prokaryotes divergence estimated by 5S RNA sequences. Nature (New Biol) 243:199–200

Lim G, Burton JC (1983) Nodulation status of the Leguminosae. In: Broughton WJ (ed) Nitrogen fixation, vol 2. *Rhizobium*. Clarendon, London, pp 1–34

McKenna MG (1975) Toward a phylogenetic classification of the Mammalia. In: Lucket WP, Szalay FS (eds) Phylogeny of the primates. Plenum Press, New York, p 21–37

Ochman H, Wilson AC (1987) Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. J Mol Evol 26:74–86

Pustell J, Kafatos FC (1982) A convenient and adaptable package of DNA sequence analysis programs. Nucleic Acids Res 10:51–59

Rossbach S, Schell J, de Bruijn FJ (1988) Cloning and analysis of *Agrobacterium tumefaciens* C58 loci involved in glutamine biosynthesis: neither the *glnA* (GSI) nor the *glnII* (GSII) gene plays a special role in virulence. Mol Gen Genet 212:38–47

Ruvkun GB, Ausubel FM (1980) Interspecies homology of nitrogenase genes. Proc Natl Acad Sci USA 77:191–195

Scott KF (1986) Conserved nodulation genes from the nonlegume symbiont *Bradyrhizobium* sp. (Parasponia). Nucleic Acids Res 14:2905–2919

Somerville JE, Kahn ML (1983) Cloning of the glutamine synthetase I gene from *Rhizobium meliloti*. J Bacteriol 156:168–176

Somerville JE, Shatters RG, Kahn ML (1989) Isolation, characterization and complementation of *Rhizobium meliloti* 104A14 mutants that lack glutamine synthetase II activity. J Bacteriol (in press)

Stadtman ER, Ginsburg A (1974) The glutamine synthetase of *Escherichia coli*: structure and control. In: Boyer PD (ed) The enzymes 10:755–807

Tingey SV, Coruzzi GM (1987) Glutamine synthetase of *Nicotiana plumbaginifolia*. Plant Physiol 84:366–373

Tingey SV, Walker EL, Coruzzi GM (1987) Glutamine synthetase genes of pea encode distinct polypeptides which are differentially expressed in leaves, roots and nodules. EMBO J 6:1–9

Tischer E, DasSarma S, Goodman HM (1986) Nucleotide sequence of an alfalfa glutamine synthetase gene. Mol Gen Genet 203:221–229

Torok I, Kondorosi A (1981) Nucleotide sequence of the *Rhizobium meliloti* nitrogenase reductase (*nifH*) gene. Nucleic Acids Res 9:5711–5723

Yang D, Oyaizu Y, Oyaizu H, Olsen GJ, Woese CR (1985) Mitochondrial origins. Proc Natl Acad Sci USA 82:4443–4447

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

Received November 16, 1988/Revised and accepted March 13, 1989