Universidad Nacional Autónoma de México

UNIDAD ACADEMICA DE LOS CICLOS PROFESIONAL Y DE POSGRADO DEL COLEGIO DE CIENCIAS Y HUMANIDADES.

CARACTERIZACION MOLECULAR DE LA NODULINA-30 DE FRIJOL.

T E S I S

Que para obtener el Grado de
DOCTOR EN INVESTIGACION BIOMEDICA BASICA
p r e s e n t a
FRANCISCO CAMPOS ALVAREZ

Departamento de Biología Molecular de Plantas del Instituto de Biotecnología.

Cuernavaca, Morelos. Mayo,1995.

UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

INDICE

		Páginas
	PRESENTACION	1
	RESUMEN	2
	ABSTRACT	3
I.	**INTRODUCCION**	4
	A. El desarrollo del nódulo	5
	B. Las Nodulinas	17
	C. Objetivos	21
	D. Antecedentes	22
	Expression of nodule-specific genes	
	in *Phaseolus vulgaris* L. (artículo)	521-532
II.	**RESULTADOS**	23
	Characterization and gene expression of nodulin	
	Npu30 from *Phaseolus vulgaris* L. (artículo)	1-24
III.	**DISCUSION Y CONCLUSIONES**	24
	A. Perspectivas	31
IV.	**REFERENCIAS**	32

ABSTRACT

We have previously reported that transcripts for a 30 kDa nodulin (Npv30) are very abundant in the nodule. This work describes the isolation and characterization of Npv30 cDNA and genomic clones. Npv30 has the following characteristic features: a) a putative signal sequence at the deduced amino terminal region; b) a proline-rich stretch at the carboxy-terminal end; and c) a characteristic domain of four cysteines resembling metal-binding sites. In *Phaseolus vulgaris* L. Npv30 is encoded by a small gene family which shares discrete sequence homologies with another small gene family in soybean. An antibody against a β-galactosidase-Npv30 fusion protein detected two proteins of 28 and 30 kDa. Although Npv30 transcripts are very abundant, they encode proteins that were hardly detected in nodule fractions, suggesting that these proteins have a short half-life and/or the mRNAs are strongly regulated at the translational level. By *in situ* hybridizations experiments, Npv30 transcripts were detected in the infected cells of the nodule.

Vo.Bo. Dr. Federico Sánchez Rodríguez
Director de Tesis.

Presentacion.

Este trabajo fué realizado en el Departamento de Biología Molecular de Plantas del Instituto de Biotecnología de la Universidad Nacional Autónoma de México bajo la dirección del Doctor Federico Sánchez Rodríguez al cual agradezco sus enseñanzas, amistad y confianza.

Los artículos que se incluyen son:

Campos F, Padilla J, Vázquez M, Ortega JL, Enríquez C and Sánchez . (1987). **Expression of nodule-specific genes in *Phaseolus vulgaris* L.** Plant Mol Biol 9: 521-532.

Campos F, Carsolio C, Kuin H, Bisseling T, Rocha-Sosa M and Sánchez F. (1995). **Characterization and gene expression of nodulin Npv30 from *Phaseolus vulgaris* L.** Plant Physiol: aceptado para su publicación.

RESUMEN.

En el presente trabajo hemos llevado a cabo la caracterización de una familia de nodulinas de 30 kDa (Npv30). La familia de la Npv30 es codificada por un grupo de transcritos que se acumulan abundantemente en los nódulos de frijol (*Phaseolus vulgaris* L.). La secuencia de aminoácidos deducida de la secuencia nucleotídica de algunos miembros de esta familia muestra que estas nodulinas poseen: 1) un posible péptido señal en el extremo amino-terminal; 2) una sección rica en prolinas en el extremo carboxilo y 3) dos regiones con el motivo Cys-X_{7AAs}-Cys, este motivo es parecido al que presentan las proteínas que unen zinc. La familia de la Npv30 es similar a la familia-A de nodulinas de soya (*Glycine max*). Aunque los niveles de la(s) proteína(s) detectada(s) por el anticuerpo contra Npv30 son muy bajos, los transcritos que codifican para esta familia son 2.5 veces más abundantes que los que codifican para la leghemoglobina, la proteína más abundante del nódulo (20% de la proteína total); esta paradoja puede ser explicada por: 1) Es posible que la(s) Npv30s tiene(n) una vida media corta, y/o 2) que haya una regulación a nivel de la traducción de los mensajeros. Por medio de hibridazión *in situ* se han detectado los transcritos de Npv30 en las células infectadas del nódulo; posiblemente el papel de la Npv30 esté relacionado con la función de estas células en el nódulo.

Vo. Bo. Dr. Federico Sánchez Rodríguez
Director de Tesis.

ABSTRACT

We have previously reported that transcripts for a 30 kDa nodulin (Npv30) are very abundant in the nodule. This work describes the isolation and characterization of Npv30 cDNA and genomic clones. Npv30 has the following characteristic features: a) a putative signal sequence at the deduced amino terminal region; b) a proline-rich stretch at the carboxy-terminal end; and c) a characteristic domain of four cysteines resembling metal-binding sites. In *Phaseolus vulgaris* L. Npv30 is encoded by a small gene family which shares discrete sequence homologies with another small gene family in soybean. An antibody against a β-galactosidase-Npv30 fusion protein detected two proteins of 28 and 30 kDa. Although Npv30 transcripts are very abundant, they encode proteins that were hardly detected in nodule fractions, suggesting that these proteins have a short half-life and/or the mRNAs are strongly regulated at the translational level. By *in situ* hybridizations experiments, Npv30 transcripts were detected in the infected cells of the nodule.

Vo.Bo. Dr. Federico Sánchez Rodríguez
Director de Tesis.

I. Introducción

Las plantas de la familia *Leguminosae* poseen la característica de poder interaccionar simbióticamente con bacterias Gram-negativas de los géneros *Rhizobium*, *Bradyrhizobium* o *Azorhizobium*. Como resultado de esta relación se forma un nuevo órgano en la raíz de la planta, el nódulo. En este nuevo órgano, la bacteria reside en el citoplasma de algunas células vegetales donde lleva a cabo la fijación de nitrógeno atmosférico por medio del sistema multienzimático de la nitrogenasa, dando como resultado la formación de amonio. Posteriormente, el amonio producido por *Rhizobium* es asimilado por el nódulo y exportado a las partes aéreas de la planta, la planta provee al microsimbionte de energía (Bauer, 1981; Brewin, 1991; Long, 1989; Sprent, 1989).

El nódulo es la culminación de una serie de eventos secuenciales que van desde el reconocimiento entre ambos simbiontes hasta la formación de un nódulo efectivo en la fijación de nitrógeno. Vincent (1980) ha agrupado el proceso de nodulación en: preinfección, infección, desarrollo y función del nódulo.

A. El Desarrollo del Nódulo

1. Preinfección

Durante esta fase hay un intercambio de señales químicas entre la bacteria y la planta, dando como resultado la

especificidad de la interacción. Los eventos que suceden en esta fase son: la colonización de la raíz por la bacteria y el reconocimiento entre la planta y *Rhizobium*, dando como indicador de una infección viable, la deformación de los pelos radiculares (p. ej. engrosamientos, enroscamientos, ramificaciones).

(a) Colonización

Se ha propuesto que la bacteria es atraída por su hospedero por quimiotaxis, de esta manera la concentración de *Rhizobium* en la rizósfera aumenta lo suficiente para que la infección sea posible (Currier y Storbel, 1979). Posiblemente los flavonoides excretados por la raíz de la planta sirven como quimioatrayentes (Caetano-Anolles *et al.*, 1988). La planta también excreta hacia la rizósfera factores que estimulan el crecimiento de *Rhizobium* sobre otros microorganismos del suelo, se cree que estos factores pueden ser sustancias tales como biotina, tiamina (Graham, 1963) u homoserina (Van Egeraat, 1972).

(b) Reconocimiento

El proceso de nodulación se inicia con el reconocimiento entre la planta y su microsimbionte, por lo menos hay dos intercambios de señales entre ambos socios; Spaink (1992) describe a este fenómeno como "preguntas y respuestas".

El primer intercambio de señales ocurre cuando la planta excreta flavonoides a la rizósfera, estos compuestos vegetales inducen la transcripción de los genes bacterianos involucrados en la nodulación (genes *nod, nol* y *syr*) (Djordevic *et al.*, 1987; Rolfe, 1980). Los flavonoides interactúan con el producto del gen bacteriano *nodD;* la proteína NodD es un factor

transcripcional que regula la activación de los operones de los genes bacterianos involucrados en la nodulación (Long, 1989).

El segundo intercambio de señales sucede cuando algunos genes bacterianos *nod* o *nol* producen las moléculas de un factor (factor Nod) identificado como un tipo lipo-oligosacárido. Dependiendo de la concentración a la que se aplique el factor Nod purificado, hay deformación de los pelos radiculares (10^{-11}M) o se estimula la división de las células corticales (10^{-7} M) de la raíz de su hospedero (Lerouge *et al.*, 1990, Roche *et al.*, 1991). El primer factor Nod identificado fue el de *R. meliloti* (NodRm-1) por Lerouge *et al.*(1990). La molécula de NodRm-1 es un tetrasacárido de D-glucosamina β-1,4 sulfatado en el extremo reductor con tres grupos aminoacetilados. En el extremo no reducido de la molécula hay un ácido graso insaturado de 16 carbonos. Se ha postulado que los productos de los genes *nodC*, *nodI*, *nodM*, *nodH*, *nodPQ*, *nodFEG* están involucrados en la síntesis del factor Nod en *R. meliloti* (ver Fig. 1)(Bruijn y Downie, 1991).

Se ha determinado que la estructura del factor Nod varía entre diferentes rhizobia y que una misma bacteria produce una gama de diferentes factores de entre 3 a 5 unidades de N-acetil glucosamina (Denarie *et al.*, 1994, Mergaert *et al.*, 1993, Price *et al.*, 1992, Sanjuan *et al.*, 1992). Cuando se aplica el factor NodRm-1 de *R. meliloti* a raíces de *Vicia sativa* no se observa ningún efecto en los pelos radiculares; sin embargo, cuando se añade el factor Nod sin el grupo sulfato, se observan deformaciones en la raíz de *Vicia*, sugiriendo que la variación

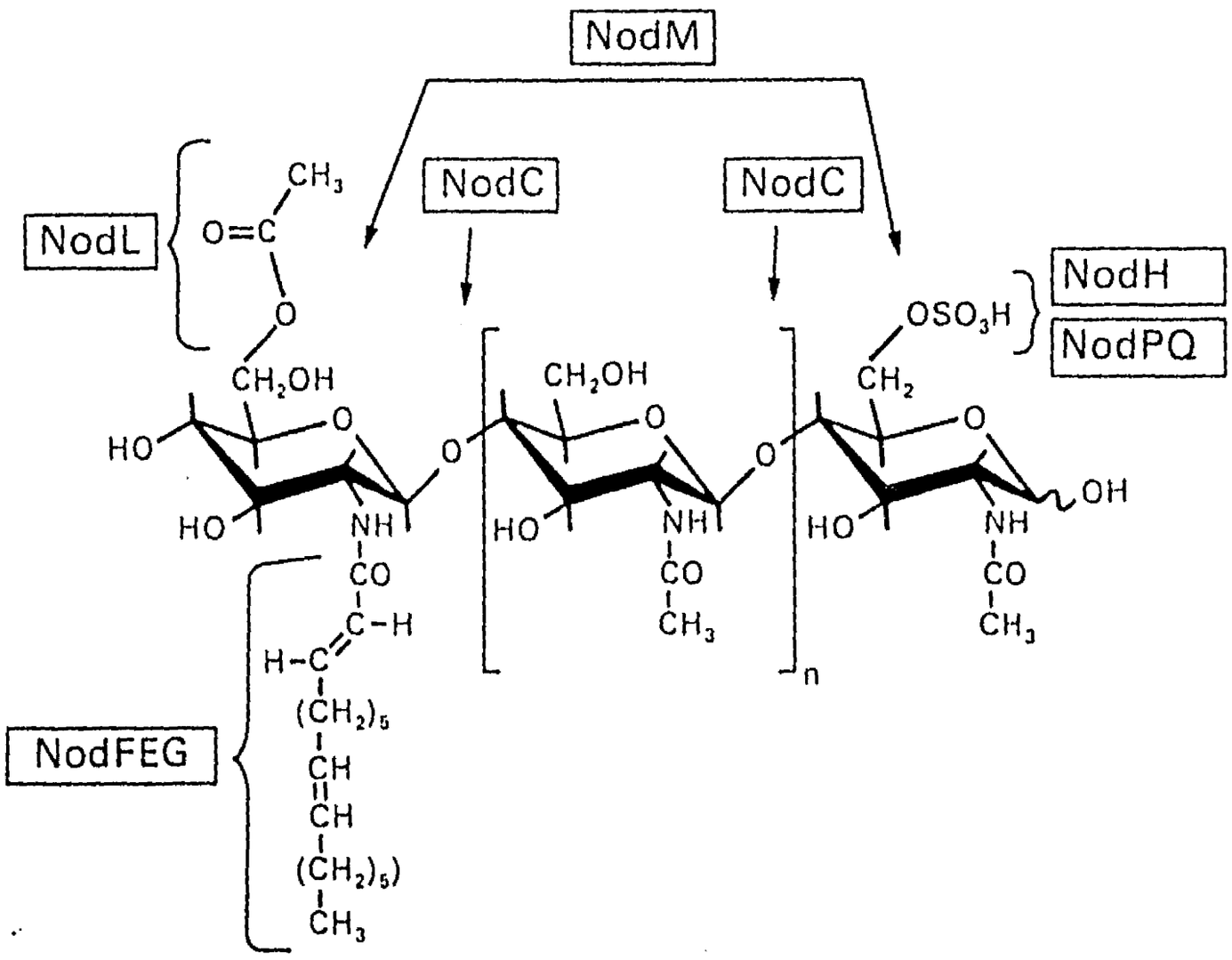

Fig. 1. **Estructura de los factores Nod.** El número de residuos de glucosamina esta indicado como **n**. En la figura se indica en que parte de la molécula estan involucrados los procutos de los genes *nod*. Tomado de Hirsch (1992).

en la estructura del factor Nod puede ser la una de las causas de la especificidad de la respuesta del hospedero (Schultze *et al.*, 1992).

Aunque los factores Nod son secretados por los rhizobia al medio, Hirsch (1992) ha propuesto que estos se encuentran anclados a la membrana de la bacteria por su parte lipídica (ver Fig. 1) y que, al hacer contacto la bacteria con el pelo radicular el factor Nod interaciona con un receptor situado en la membrana del pelo radicular. La activación del receptor por los factores Nod desencadenaría una serie de cambios subcelulares en el pelo radicular como son: el cambio de permeabilidad de la membrana y un rearreglo del citoesqueleto. Long y Cooper (1988) han propuesto que estos cambios son o generan "una segunda señal" para la inducción de la división de las células corticales. Otra posibilidad, es la existencia de dos ó más receptores con diferentes afinidades para los diferentes factores Nod, tanto para bajas concentraciones (10^{-11}; deformación de pelos radiculares) como para altas concentracionens (10^{-7}; división de células corticales) (Hirsch, 1992).

2. Infección

En la mayoria de las leguminosas, durante el proceso de infección los pelos radiculares son penetrados por la bacteria,

a través de una estructura tubular llamada hilo de infección (Robertson y Farnden; 1980).

Como preludio a la formación del hilo de infección se observa el acercamiento del núcleo de la célula radicular hacia el sitio de la infección (Napoli y Hubell, 1975).

La infección comienza con una hidrólisis localizada de la pared celular del pelo radicular. El mecanismo de la formación de esta "lesión" no esta bien elucidado. Verma *et al.*, (1978a) han detectado histoquímicamente actividad celulolítica en las paredes del hilo de infección de soya. Se ha propuesto la acción de enzimas pectolíticas o celulolíticas de *Rhizobium* para la formación del hilo de infección (Fahraeus y Ljunggren, 1959; Ljunggren y Fahraeus, 1961). En el sitio de la "lesión" de la pared celular, la bacteria penetra al pelo radicular por una invaginación de la membrana y la planta formará una estructura tubular alrededor de la bacteria; dicha estructura está hecha de material parecido al de la pared celular. En este tubo la bacteria se encuentra en una matriz de mucopolisacáridos; conforme el hilo de infección penetra más capas de células de la raíz; la bacteria prolifera dentro de este tubo (aproximadamente tres días después de la inoculación). Cuando el hilo de infección atraviesa de 3 a 6 capas de células corticales externas, se inicia una zona de actividad meristemática, la cual es el blanco de la infección. Las células de esta zona, principalmente poliploides, serán infectadas por *Rhizobium*. El resto de las células de la zona meristemática consiste de células diploides, las cuales darán lugar a las células no

infectadas en el nódulo maduro (Newcomb, 1981; Robertson y Farden, 1980).

El *Rhizobium* es liberado al final del hilo de infección al citoplasma de las células meristemáticas poliploides por un mecanismo de fagocitosis (Robertson y Farden, 1980). Las bacterias liberadas son envueltas en membrana de origen vegetal, denominada membrana peribacteroidal (Robertson *et al.*, 1978a; Robertson *et al.*, 1978b; Verma *et al.*, 1978b).

3. Desarrollo del Nódulo

Los nódulos de las plantas leguminosas se encuentran agrupados, según su forma, en: esféricos, con un crecimiento determinado, como los de frijol y soya; apicales o cilíndricos, con un crecimiento no determinado o apical, como los de trébol y alfalfa (Allen y Allen, 1981) (ver Fig. 2).

La forma del nódulo es el resultado del desarrollo de la actividad mitótica (Newcomb, 1981), en el caso de los nódulos esféricos o globosos la actividad mitótica se lleva a cabo alrededor de la zona meristemática original, a diferencia de lo que sucede en los nódulos cilíndricos donde la actividad se localiza en el ápice. En los nódulos esféricos, después del vaciamiento de los rhizobia, las células del hospedero cesan su mitosis y sufren un aumento de volumen. Las células infectadas de la zona central causan que el nódulo crezca radialmente, lo que hace que el nódulo adquiera una forma esférica (Newcomb, 1981; Robertson y Farden, 1980).

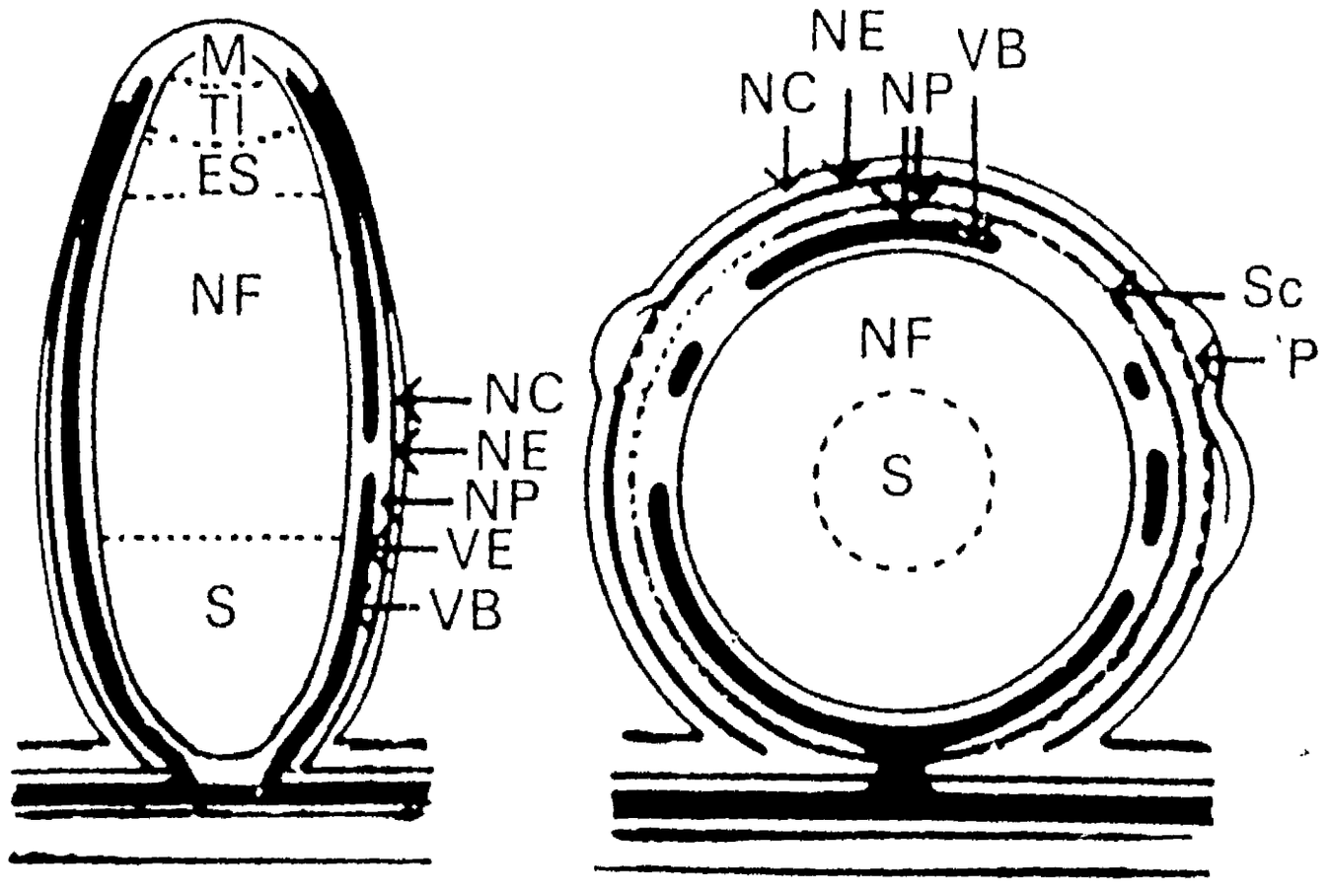

Fig.2. **Comparación entre nódulos esféricos y cilíndricos.** En esta figura se muestran las diferencias entre nódulos cilíndricos y esféricos. Abreviaciones: M, meristemo; TI, zona de hilos de infección; ES, zona simbiótica temprana; NF, zona de fijación de nitrógeno; S, zona senecente; NC, corteza del nódulo; NE, endodermis del nódulo; NP, parénquima del nódulo; VE, endodermis vascular; VB, haces vasculares; Sc, esclerénquima; P, peridermo. Tomado de Hirsch, (1992).

(a) Histología del Nódulo

El nódulo está compuesto por células infectadas y no infectadas. Las células no infectadas son más pequeñas que las infectadas. Otro tipo de células asociadas al nódulo son las que forman los haces vasculares y las células de transferencia, estas células están relacionadas al transporte de metabolitos (Newcomb, 1981).

4. Función del Nódulo

Poco después de que los rhizobia se encuentran en el citoplasma del hospedero y rodeado por la membrana peribacteroidal, éste se diferencia a bacteroide, la forma especializada de la bacteria que reduce nitrógeno atmósferico a amonio. El amonio es excretado al citoplasma de la célula infectada, el cual es asimilado por el hospedero en el nódulo y finalmente metabolizado en compuestos nitrogenados que serán exportados para nutrir a las partes aéreas de la planta. Por otro lado, la planta contribuye proporcionando un medio adecuado (niveles de O_2 libre muy bajos) y una fuente de energía que le permite al bacteroide fijar nitrógeno (Robertson y Farnden, 1980).

(a) Fijación de Nitrógeno

Un resultado fisiológico de la simbiosis es la desrepresión de la dinitrogenasa. La diferenciación de las células del nódulo

puede ser vista como una estrategia para optimizar este proceso (Werner, 1992).

La enzima responsable de la fijación de nitrógeno es la nitrogenasa (dinitrógeno oxidorreductasa: hidrolizadora de ATP: EC 1.18.6.1), la cual cataliza la siguiente reacción:

$$N_2 + 8\ e^- + 8\ H^+ + 16\ MgATP \longrightarrow 2NH_3 + H_2 + 16MgADP + 16P_i$$

La dinitrogenasa está formada por dos componentes, una proteína que contiene fierro y molibdeno de 245 kDa y una ferroproteína de 64 kDa. La Mo-Fe-proteína es un tetrámero que consiste de dos subunidades distintas de 51.3 y 50 kD($\alpha\alpha\beta\beta$). La Fe-proteína consta de dos subunidades idénticas de 34.6 kD. Ambos componentes son desnaturalizados rápida e irreversiblemente por exposición a O_2. En la nueva nomenclatura la Mo-Fe-proteína es llamada dinitrogenasa y la Fe-proteína, dinitrogenasa reductasa (Werner, 1992).

(b) Metabolismo del Carbono y del Nitrógeno en el Nódulo

Como consecuencia de la simbiosis, el hospedero tiene que reestructurar las vías metabólicas del carbono de la planta para proveer de fuentes de carbono tanto para el bacteroide como para poder asimilar y translocar el NH_4^+ fijado .

(c) Asimilación del Amonio

Aproximadamente el 15% del nitrógeno fijado por el bacteroide es exportado como amonio al citoplasma de la célula

infectada (Lea *et al.*, 1982). Ohyma y Kumazawa (1980) usando inhibidores específicos de la glutamino sintetasa (GS) y glutamato sintasa (GOGAT), tales como metionina sulfoximina y azaserina, han establecido que el nitrógeno fijado se asimila por la vía GS/GOGAT.

Robertson *et al.*, (1975) han reportado que la actividad de GS en nódulos de *Lupinus* aumenta conforme transcurre la nodulación, llegando a ser 500 veces más alta que la que se detecta en la raíz. La GS está formada por un octámero de 380 kD que consiste de 8 subunidades de 43 y 47 kD y constituye el 2% de la proteína soluble del nódulo en soya y frijol (McParland *et al.*, 1976; Lara *et al.*, 1984). Esta enzima es citosólica (Awoniake *et al.*, 1981). Es posible separar dos formas de GS en los nódulos de frijol (GSn_1 y GSn_2) (Cullimore *et al.*, 1983). GSn_1 es una forma específicamente asociada al nódulo, mientras la forma GSn_2 es similar a la de raíz (Lara y cols., 1983). Del análisis de la composición polipeptídica de GSn_1, Lara *et al.*, (1983) encontraron que consta de dos isoformas denominadas β y γ, siendo el péptido γuna proteína específica del nódulo.

La enzima GOGAT consiste de un solo péptido de 235 kD (Boland y Benny, 1977), en *Lupinus* la actividad aumenta hasta 10 veces durante la nodulación (Robertson *et al.*, 1975) sin embargo, en los nódulos de soya no se observa un aumento en la actividad (Sen y Schulman, 1980), en frijol, en cambio si hay un aumento correlacionado con la fijación del nitrógeno (Lea *et al.*, 1982).

(d). Transporte del Nitrógeno Asimilado.

El nitrógeno asimilado se exporta del nódulo por el xilema o floema al resto de la planta en forma de asparagina, en el caso de leguminosas "templadas" como *Lupinus*, chícharo y alfalfa. Esta amida constituye el 90% del nitrógeno transportado del nódulo hacia las hojas y semillas en desarrollo (Pates *et al.*, 1969; Atkins *et al.*, 1975; Robertson *et al.*, 1975). El N_2 es incorporado hacia asparagina en una vía que involucra dos enzimas: aspartato aminotransferasa y asparagina sintetasa (Lea y Miflin, 1980; Miflin y Cullimore, 1984). En cambio en las leguminosas "tropicales" como soya, frijol y *Vigna* el nitrógeno es transportado como alantoína y ácido alantoico (ver Fig. 3) (Reynolds *et al.*, 1982).

Aunque en plantas y animales los ureidos son derivados de la degradación de purinas, en el nódulo la síntesis de ureidos está separada del metabolismo normal de ácidos nucleicos (Fig.3) (Lea *et al.*, 1982): algunas enzimas de esta vía son: xantina deshidrogenasa, uricasa y alantoinasa, las cuales aumentan su actividad cuando se incrementa el transporte de ureidos en el nódulo (Schubert, 1981). La xantina deshidrogenasa ha sido purificada por Triplet *et al.*(1981) y consta de un dímero de 285 kD. La uricasa presente en nódulos de soya y frijol es diferente a la que se encuentra en raíz y consta de 4 subunidades de 35 kD (Bergman *et al.*, 1983; Sánchez *et al.*, 1987).

Se ha propuesto que las células no infectadas del nódulo participan en la síntesis de ureidos (Newcomb y Tandon, 1981) debido a que estas células tienen más peroxisomas y retículo

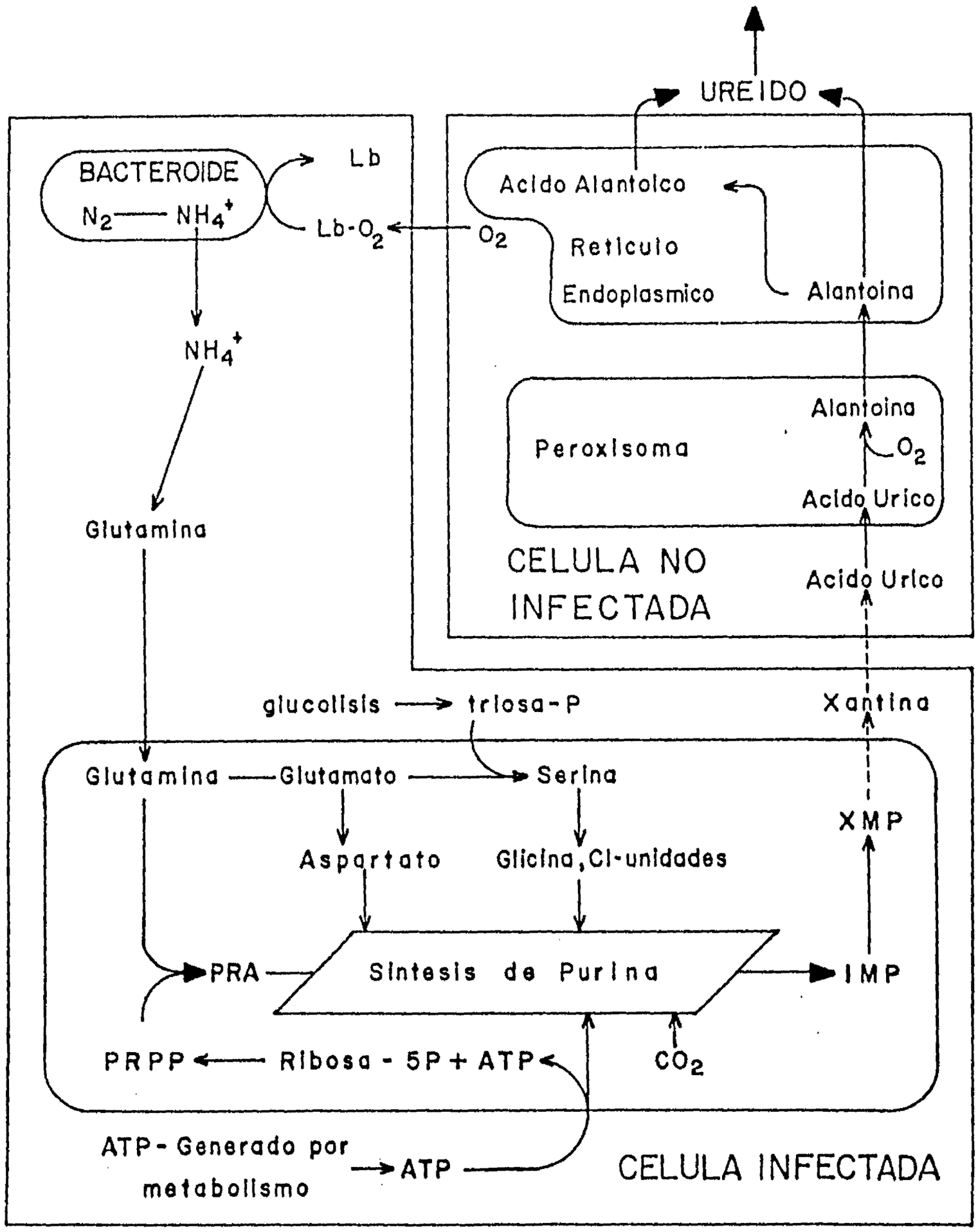

Fig. 3. **Metabolismo del carbono y nitrógeno en células infectadas y no infectadas de nódulos de leguminosa.**

Tomado de Dilworth y Glenn, 1984.

endoplásmico liso, lugares donde se localiza preferencialmente la uricasa y la alantoinasa (Hanks *et al.*, 1981; Hanks *et al.*, 1983). Bergman *et al.* (1983) han demostrado que la uricasa del nódulo de soya se encuentra en peroxisomas de células no infectadas.

B. Las Nodulinas.

Para que se lleve a cabo la formación de un nódulo efectivo en fijación del nitrógeno se necesita de la expresión diferencial y coordinada de los genomas de la planta y la bacteria.

El estudio de los genes que se expresan durante la formación del nódulo se encuentra mucho mas adelantado en la bacteria, que en la planta (Kondorosi, 1990, Kolchinsky *et al.*, 1994); esto se debe principalmente a que en la actualidad está mas avanzada la aplicación de las técnicas de Biología Molecular en bacterias Gram-negativas que en plantas (p. ej., obtención de mutantes, complementación, transferencia de DNA, etc.). En esta sección revisaremos los genes de la planta que se encuentran involucrados en la simbiosis.

En la planta se han identificado pocos genes involucrados en la simbiosis con la bacteria y la función de la mayoría de ellos es deconocida (Sánchez *et al.*, 1991).

Al comparar los productos de traducción *in vitro* de RNA poli(A^+) de raíz y nódulo de soya, Legocki y Verma (1980) han descrito la presencia de proteínas de la planta, específicas del nódulo en soya, las cuales denominaron nodulinas. Se han encontrado nodulinas en chícharo (Bisseling *et al.*, 1983),

alfalfa (Lang-Unnasch y Ausubel, 1985) y frijol (Campos *et al.*, 1987). Usando esta estrategia, se han detectado 20 nodulinas en soya (Legocki y Verma, 1979; Legocki y Verma, 1980), 21 nodulinas en chícharo (Govers *et al.*, 1985) y alrededor de 20 nodulinas en frijol (Campos *et al.*, 1987). Otro enfoque usado para la determinación y aislamiento de genes específicos del nódulo, es la detección de clonas de cDNA que codifican para nodulinas por medio de un tamizado diferencial en bancos de cDNA de nódulo (Govers *et al.*, 1987; Campos *et al.*, 1987; Fuller *et al.*, 1983).

Basándose en su tiempo de aparición, las nodulinas han sido clasificadas en nodulinas tempranas y nodulinas tardías (Nap y Bisseling, 1990). Las nodulinas tempranas son detectadas durante la preinfección, infección y formación del nódulo. Para la inducción de nodulinas tempranas no es necesaria la presencia de la bacteria ya que la sola adición del factor Nod a la raíz es capaz de inducir la expresión de estos genes (Sánchez *et al.*, 1991). También se ha detectado la expresion de estas nodulinas en "pseudonodulos" inducidos con inhibidores del transporte de auxina en raices de alfalfa y en estructuras parecidas a nódulos las cuales se forman espontaneamente en raices de alfalfa (Hirsch, 1992)

La mayoría de las nodulinas tempranas han sido aisladas por el grupo de Ton Bisseling. Se han caracterizado en soya (GmENOD = nodulina temprana de *Glycine max*) y chícharo (PsENOD = nodulina temprana de *Pisum sativum*) alrededor de 10 nodulinas tempranas (Bisseling *et al.*, 1990). La mayoría son proteínas

ricas en prolina y con repetidos de aminoácidos similares a los encontrados en proteínas asociadas a la pared celular (PsENOD2, PsENOD12, GmENOD13, GmENOD55) (Bisseling *et al.*, 1990). PsENOD5 tiene similitud con arabinogalactanas y se ha detectado su trancrito en células de alrededor del hilo de infección (Scheres *et al.*, 1990). Se ha determinado que la nodulina temprana ENOD12 puede ser inducida cuando se tratan las raíces de chícharo con el factor Nod de *R. leguminosarum bv. viciae*, lo que indica que la expresión de estos genes puede estar controlada por el microsimbionte (Horvath *et al.*, 1993).

Se ha encontrado una nodulina temprana en soya y alfalfa denominada ENOD40 (Yan *et al.*, 1993; Kouchi y Hata, 1993); aunque el transcrito de ENOD40 esta poliadenilado no codifica para alguna proteína. Kondorosi *et al.*, (1994) han determinado que los transcritos de ENOD40 tienen un alto grado de estabilidad, una caracteristica de los RNAs no codificantes y que la sobreexpresión de ENOD40 en embriones de *Medicago* causa que ellos se desarrollen como teratomas, estos datos sugieren que posiblemente el transcrito de ENOD40 sea un "riborregulador" y que su expresión tenga algún papel en el desarrollo del nódulo, ya que estos RNAs estan asociados con el crecimiento y la diferenciación célular.

Las nodulinas tardías son genes que se expresan un poco antes de que la fijación de nitrógeno se inicie y están involucradas en la función del nódulo. Fuller *et al.*, (1983) han clasificado las nodulinas en tres clases, basándose en su posible función : (1) Proteínas responsables para el

mantenimiento del nódulo; (2) enzimas necesarias para la asimilación específica del nitrógeno reducido; y (3) proteínas que apoyan la función del bacteroide y facilitan la reducción del nitrógeno. Se conoce la función de algunas proteínas específicas del nódulo, de entre ellas sólo una proteína del tercer tipo: la leghemoglobina (Verma y Bal., 1976; Appleby, 1984); dos del segundo: una forma de uricasa específica del nódulo de soya y frijol (Bergmann *et al.*, 1983; Sánchez *et al.*, 1987), el polipéptido γ de la enzima glutamino sintetasa en nódulos de frijol (Lara *et al.*, 1983; Ortega *et al.*, 1986; Padilla *et al.*, 1987); y dos de la primera categoría,una sacarosa sintetasa nódulo-específica (Thummler y Verma, 1986) y una colín-cinasa específica de la membrana peribacteroidal en soya (Mellor *et al.*, 1986). La función del resto de la nodulinas tardías es desconocida.

Se han detectado actividades enzimáticas elevadas de enzimas de la vía de los ureidos en el nódulo, como son: xantina deshidrogenasa, fosfoenol-piruvato carboxilasa, purina nucleosidasa y malato dehidrogenasa (Robertson y Farnden, 1980), es probable que existan isoformas específicas del nódulo para algunas de estas enzimas.

Aunque Legocki y Verma (1980) junto con van Kammen (1984) han definido como nodulinas a aquellas proteínas que se expresan sólo en nódulo y no en la raíz sin infectar por *Rhizobium;* no se ha establecido en que estadío del desarrollo de la raíz debe llevarse a cabo esta comparación, lo que ha causado que en la mayoria de los reportes donde se describen nodulinas, sólo se

use como "raíz" material sin inocular de 7 días, obviando que algunas nodulinas podrian ser expresadas en la raíz a lo largo de su desarrollo. Lo más adecuado seria que junto con las cinéticas de crecimiento del nódulo se incluyeran los mismos tiempos de desarrollo de raices no infectadas, sin embargo, la solución a este problema no es tan facil, ya que a una planta sin inocular se le deben añadir compuestos nitrogenados lo cual seria una variable más para estos experimentos. A pesar de que no se ha estudiado más a fondo este problema, los datos obtenidos en experimentos de hibridazión *in situ* e inmunolocalización de algunas nodulinas han confirmado la definición de "nodulina", ya que en estos el análisis se lleva a cabo junto con la raíz a la que esta unido el nódulo (Scheres et al., 1990;).

C. Objetivos

Las finalidades del presente trabajo son dos: (1) identificar los genes de la planta que están involucrados en la simbiosis frijol-*Rhizobium*, y (2) determinar la función de estos genes durante el proceso simbiótico. Durante mi trabajo de Maestría y parte del de Doctorado abordé el primer punto, culminando en la identificación de alrededor de 20 nodulinas en frijol (ver Antecedentes).

Un hallazgo muy interesante al describir la presencia de nodulinas tardías en frijol, fue la detección de la nodulina-30 o bien, Npv30, usando la nomenclatura propuesta por van Kammen (1984). La Npv30 posee una serie de características muy particulares como son: (a) ser codificada por transcrito(s) muy

abundante(s) en el nódulo, y (b) aparentemente estar codificada por una familia multigénica.

El objetivo principal de mi proyecto de Doctorado fue el de llevar a cabo una caracterización molecular mas exhaustiva de la Npv30, con la finalidad de poder determinar su posible papel en el nódulo.

Objetivos específicos de este trabajo.

1. Determinación y caracterización de la secuencia nucleotídica del gene que codifica para la Npv30.

2. Cuantificación de los niveles del transcrito en el nódulo.

3. Determinar si la Npv30 es codificada por una familia de genes.

4. Localización celular de la nodulina-30.

D. Antecedentes

En la siguiente sección de esta Introducción presento el artículo "**Expression of nodule-specific genes in *Phaseolus vulgaris* L.**" publicado en la revista Plant Molecular Biology 9: 521-5332. Este artículo es el producto de mi trabajo de Maestría y parte del de Doctorado.

Plant Molecular Biology 9: 521–532 (1987)
© Martinus Nijhoff Publishers, Dordrecht – Printed in the Netherlands

Expression of nodule-specific genes in *Phaseolus vulgaris* L.*

Francisco Campos, Jaime Padilla, Martha Vázquez, José Luis Ortega, Consuelo Enríquez and Frederico Sánchez
Departamento de Biología Molecular de Plantas, Centro de Investigación sobre Fijación de Nitrógeno, U.N.A.M., Apartado Postal 565-A, Cuernavaca, Morelos, Mexico

Received 17 November 1986; accepted in revised form 21 July 1987

Key words: gene expression, nitrogen fixation, nodulins, *Phaseolus vulgaris*, *Rhizobium*, root-nodule development

Abstract

The identification of some nodule-specific host proteins (nodulins) from common bean (*Phaseolus vulgaris* L.), a tropical ureide-transporting legume, is described. Particularly, the existence and developmental expression of several abundant nodule-specific transcripts of *P. vulgaris* are shown, including leghemoglobin, nodule-specific uricase and a group that *in vitro* translates into a cluster of about 30 kDa products. The expression pattern of nodulins in effective (Fix^+) nodules compared to ineffective (Fix^-) ones is also presented. The modified expression of main nodulins observed between these nodules indicates that different levels and/or factors associated with their regulation are involved. The intracellular infection by *Rhizobium* as a decisive step in the induction of some *P. vulgaris* nodulins is discussed.

Introduction

During the symbiotic interaction between soil bacteria of the genus *Rhizobium* and roots of its host-legume, nitrogen-fixing nodules are induced. An effective root-nodule is the culmination of a series of complex development events which includes the root-hair curling, the formation of an infection thread, the induction of meristematic activity and infection of cortical root cells, as well as the differentiation of bacteria into nitrogen-fixing bacteroids [23, 35] and the induction of the host metabolic machinery for ammonia assimilation, biosynthesis and distribution of nitrogenous compounds [28]. Nodulated temperate legumes (e.g., alfalfa, peas) transport amides in the xylem, in contrast to tropical ones (e.g., soybeans, common beans and cowpeas) which preferentially export ureides from nodules [2]. Increased activity levels of enzymes involved in the formation of ureides have been detected in soybean nodules [27].

The current approach followed to settle the spatial and temporal symbiotic response of nodulating legumes has been to study the expression of nodule-specific proteins (nodulins) and their corresponding transcripts during development. The presence of some nodulins from soybean [6, 7, 18], pea [8], and alfalfa [13] nodules, has been reported. Important functions concerned with the structure and metabolism have been inferred from the abundance and organ-specificity of nodulins [33]. However, except

*Supported in part by grants from Consejo Nacional de Ciencia y Tecnología (CONACyT), grant PCCBBNA-022632; from Fondo de Estudios e Investigaciones Ricardo J. Zevada and from the US National Academy of Sciences/National Research Council by means of a grant from the US Agency for International Development.

for the leghemoglobins [17], nodule-specific uricase II [3], sucrose synthetase [32], a peribacteroidal membrane-bound choline kinase in soybean [22], and the nodule-specific glutamine synthetase (GS-gamma) in *P. vulgaris* [14], the location and function of many other nodulins remain unknown. The identification of nodule-specific gene products will help to establish the role and requirements of the host-plant for the specificity and effectiveness of the symbiotic response.

In this work, *Phaseolus vulgaris* nodulins were identified by electrophoretic analysis of the *in vivo* and *in vitro* synthesized proteins. Recombinant plasmids encoding for these products were isolated from a cDNA library constructed from nodule polysomal RNA. The developmental expression pattern of some abundant nodulins suggested different stages and mechanisms for nodulin induction. Some of these nodulins will be useful phenotypical markers for analyzing the time-course of the bacteria-plant interaction in defective symbiosis.

Material and methods

Biological materials

Bean (*Phaseolus vulgaris* L., cv. Negro Jamapa) seeds were germinated for two days and transferred to pots, inoculated with a rifampicin-resistant derivative of *Rhizobium phaseoli* strain CIAT899 [10], or with the broad host range *Rhizobium* sp. strain CFN400 (=Cli80) isolated from nodules of the legume *Clitoria ternantea* [21]. Plants were grown as described previously [14]. Root-nodules and uninoculated roots were harvested at different times as indicated and stored at −70 °C until use.

Nodule fractionation and protein extraction

Root and nodule soluble protein fractions were prepared by grinding approximately 10 g of frozen tissue with solid CO_2 and by homogenizing at 4 °C in 10 ml of a 50 mM Tris-HCl pH 7.5 extraction buffer containing 5 mM 2-mercaptoethanol (2-ME), 1 mM phenylmethylsulphonate-fluoride and 5% w/w polyvinyl polypyrrolidone. The homogenate was filtered through cheesecloth and centrifuged at 10000 × *g* for 15 min. This supernatant is referred to as the soluble protein fraction. The isolation of an enriched peribacteroidal (membrane plus fluid) fraction and bacteroids was carried out as previously reported [1]. Protein was measured as described by Lowry *et al.* [19].

Gel electrophoresis of proteins

Proteins were separated on 12 or 7.5–15% linear gradient SDS-polyacrylamide gel electrophoresis (SDS-PAGE) according to Laemmli [12]. Two-dimensional gel electrophoresis (2D-PAGE) was carried out as described by O'Farrell [24]. Gels were stained with Coomassie blue R-250 and those with [^{35}S]-labeled proteins were processed for fluorography according to Laskey & Mills [16].

Polysomal RNA isolation and Northern blot analysis

Polyribosomes were isolated from nodules and roots following the procedure reported by Jackson & Larkins [11]. Total polysomal RNA was extracted from polyribosomes as described by Christofferson & Laties [5], omitting the proteinase-K step. Northern blots were prepared by electrophoresis of 10 μg polysomal RNA in agarose gels containing 2.2 M formaldehyde according to Maniatis *et al.* [20], transferred onto nitrocellulose paper [31] and hybridized with [^{32}P]-labeled inserts of nodule-specific clones.

Hybrid selection and in vitro *translation*

Hybrid selection was done according to the procedure described by Fuller *et al.* [6], except that filters were hybridized with 150 μg of polysomal RNA. Total polysomal RNA (10 μg) or hybrid-selected RNA were translated *in vitro* in 25 μl of rabbit reticulocyte lysate [26] containing 50 μCi of [^{35}S]-methionine (1255 Ci/mmol, Amersham) and incubated for 90 min at 30 °C.

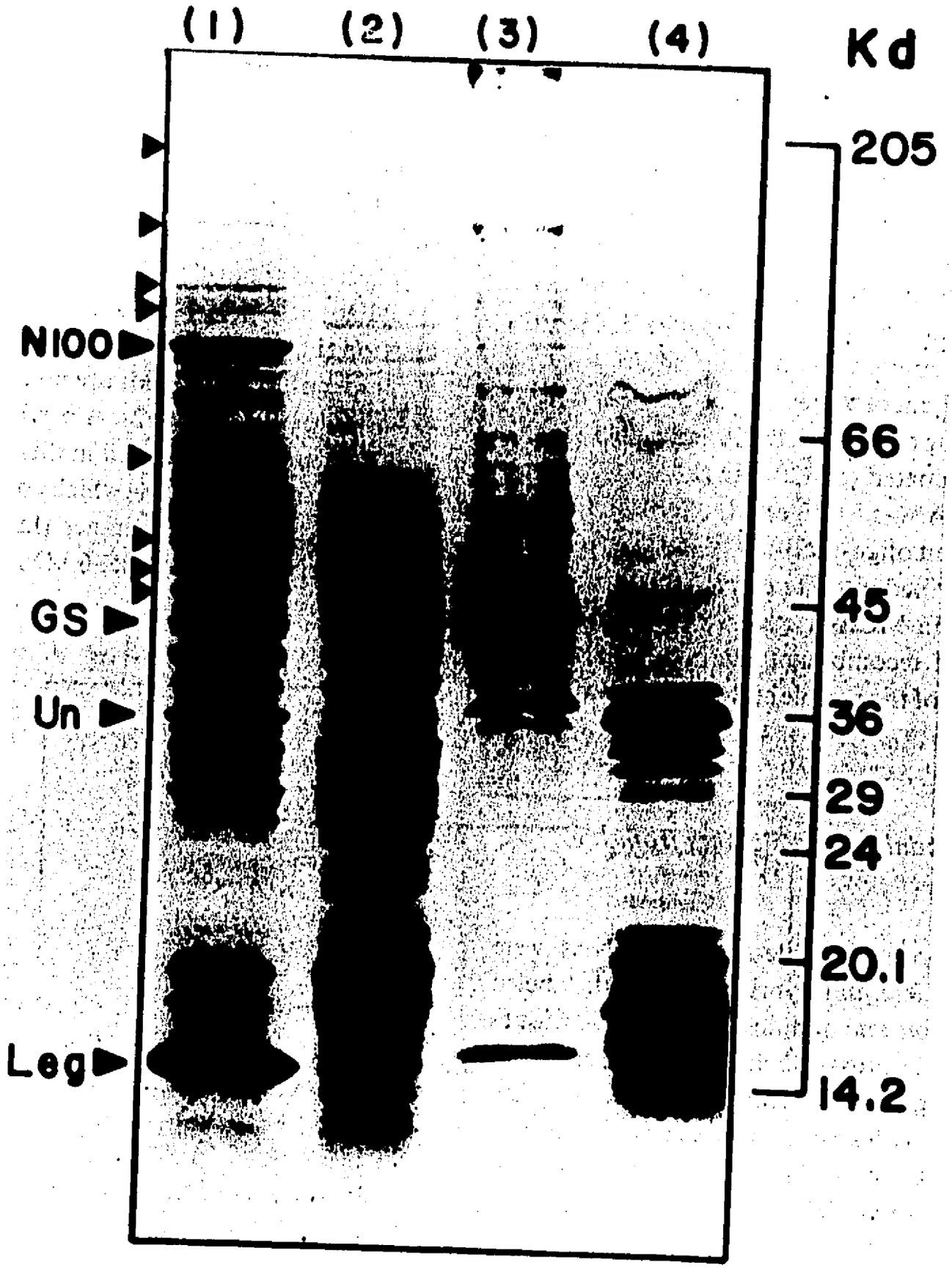

Fig. 1. Nodule-specific proteins from *P. vulgaris*. 150 μg of protein from: (1), 18-day nodule soluble fraction; (2), 8-day root soluble fraction; (3), peribacteroidal fraction and (4), bacteroid lysate were subjected to 7.5–15% SDS-PAGE and stained with Coomassie blue. Positions of nodulin-100 (N-100), glutamine synthetase (GS), uricase II (Un), and leghemoglobin (Leg) are indicated. Other putative nodulins in the 50–200 kDa range are pointed by smaller arrowheads.

Construction of a cDNA library from P. vulgaris *nodule poly(A⁺) RNA*

Nodule poly(A⁺) RNA was isolated from total polysomal RNA of 26-day nodules as previously described [20]. 5 μg of poly(A⁺)RNA were denatured in 10 mM methylmercuric hydroxide and used to prepare a cDNA library in pBR322. The first strand was synthesized with Reverse Transcriptase (Life Sciences), at a concentration of 20 units/μg of RNA. The final concentration of RNA in the mixture was 20 μg/ml in a 100 mM Tris-HCl pH 8.3 reaction buffer containing 10 mM $MgCl_2$, 100 mM of dATP, dGTP, dTTP and 0.5 mM dCTP and

100 μCi [^{32}P]-dCTP. It also included RNAsin (Promega Biotech) at 10 units/μg of RNA, 100 μg/ml of oligo-(dT)12–18 (PL Biochemicals) and 20 mM 2-ME. Reaction was brought to 20 mM EDTA and chromatographed through Sephadex G-75; the yield was approximately 70%. Second-strand synthesis was initiated by boiling for 3 min and immediately quenching RNA-cDNA hybrids on ice. The reaction mixture of 100 mM K-HEPES pH 6.9 buffer containing 100 units of *Escherichia coli* DNA polymerase per μg of single-stranded cDNA, 5–20 ng of nodule cDNA and 1 mM of each dNTP was incubated for 120 min at 15 °C. Double-stranded cDNA was S1-digested, oligo-(dC)-tailed, and cloned in an oligo-(dG)-tailed *Pst*I-cut pBR322 vector as described elsewhere [20]. Transformation of *E. coli* strain MC1061 gave 1×10^6 transformed cells per μg of recombinant plasmid. The insert's average length of the nodule cDNA library was nearly 700 bp.

Isolation of nodule-specific clones from the cDNA library

Colonies harboring nodule-specific cDNAs were selected by differential hybridization with labeled cDNA from roots and nodules of *P. vulgaris*. Clones were examined by cross-hybridization with heterologous (leghemoglobin and nodulin-24 from soybean) probes [6, 7] or with an homologous probe of uricase II (pNF-UR07), obtained from an expression cDNA library of *P. vulgaris* made in pUC9 [29].

Results

Electrophoretic analysis of soluble protein fractions from root and nodule

As an initial approach to identify nodule-specific proteins in *Phaseolus vulgaris*, the soluble protein pattern from root, nodule, and enriched peribacteroidal and bacteroidal fractions were compared (Fig. 1). Several protein bands were apparently only found in the nodule soluble and peribacteroidal fractions. The 14 kDa protein was the most abundant and corresponds to leghemoglobin; this protein band was resolved in 4 isoforms in 2D-PAGE (unpublished). The 34 and 100 kDa bands were next in abundance. The first one corresponds to uricase II and represented 2% of the total nodule soluble protein in *P. vulgaris* [29]. The 100 kDa band could be equivalent to soybean sucrose synthetase based on data recently reported [32]. There are at least 8 additional specific protein bands in the nodule soluble fraction (Fig. 1, lane 1). In the root soluble fraction, a number of protein bands are no longer detected in the nodule (Fig. 1, lane 2). A peribacteroidal membrane fraction was included in this analysis in order to identify those nodulins which could be associated to peribacteroidal membranes (lane 3). An abundant protein band of about 36 kDa, two bands in the

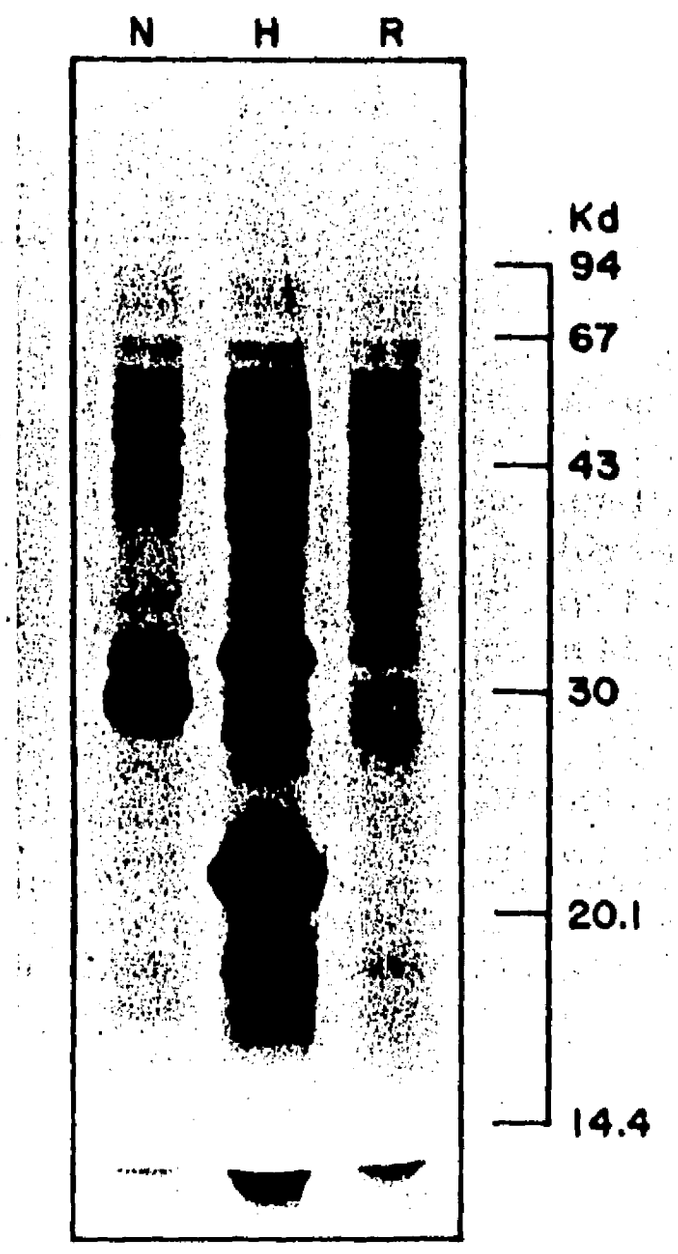

Fig. 2. Nodule-specific gene products from *in vitro* translation of *P. vulgaris* RNA. Polysomal RNA extracted from 20-day nodules (N), trifoliate leaves (H) and uninoculated 8-day roots (R) were translated and 5×10^5 cpm loaded per lane, subjected to 12% SDS-PAGE and fluorographed. MW markers are indicated.

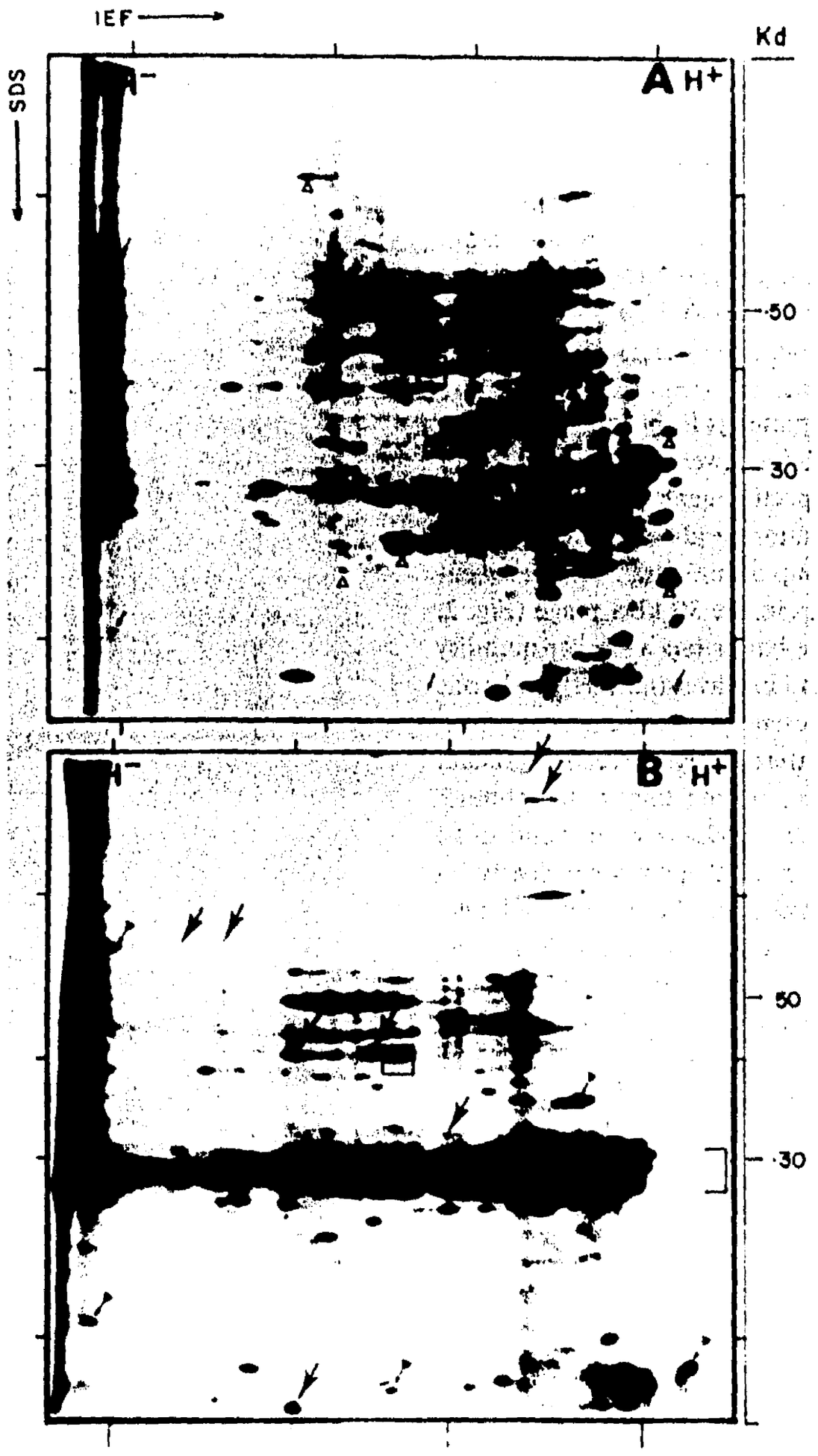

Fig. 3. Two-dimensional pattern of *in vitro* translation products from (A), uninoculated root and (B), 18d nodule polysomal RNA. Symbols indicate: open triangles, root-specific products; large arrows, nodule-specific products. Products present in root, small arrows, which were relatively induced in nodules, small arrows with triangles. The 30 kDa nodulin cluster is indicated ( ]); the GS-beta product (enclosed in a square) was used as position marker.

vicinity of 60 kDa and some bands in the 150–200 kDa range were found to be enriched in this fraction. The 14 kDa band could be a contaminant of leghemoglobin (lane 3). Major protein bands detected in the nodule soluble and peribacteroidal fractions were not derived from bacteroids (lane 4).

Electrophoretic analysis of products from in vitro *translation of polysomal mRNA from root and nodule*

SDS- and 2D-PAGE patterns from root and nodule *in vitro* translated products were compared in order to identify nodule-specific gene products (Figs. 2 and 3). *In vitro* translation products from *P. vulgaris* nodules showed a group of main transcripts that encode for polypeptides in the 30 kDa range (Fig. 2, lane N). One of these bands has a similar mobility to a polypeptide found in leaves (lane H), and none of them seem to be present in roots (lane R). By two-dimensional (2-D) pattern analysis, we found that at least 10 polypeptides were unique to nodules (Fig. 3). About 5 other products were found to be common to roots and nodules but were clearly increased in the nodule pattern. The group of polypeptides in the 30 kDa range displayed a peculiar electrofocussing pattern which appeared as a 'streak' across the IEF gradient (Fig. 3B). These results suggest that the 30 kDa nodulin cluster could be derived from multiple related transcripts.

cDNA cloning and identification of the mRNA encoding for main nodulin transcripts

About 20% nodule-specific clones were isolated from the cDNA library by differential hybridization with root cDNA. Approximately 10% of these clones hybridized to a soybean leghemoglobin probe. A homologous *P. vulgaris* clone (pNF-Lb01) was isolated (not shown). A number of highly expressed clones (around 20%) were selected. One of them (pNF-N30-1, 700 bp insert) was found to hybrid-select an mRNA that encoded for an *in vitro* translation product which migrated as one of the

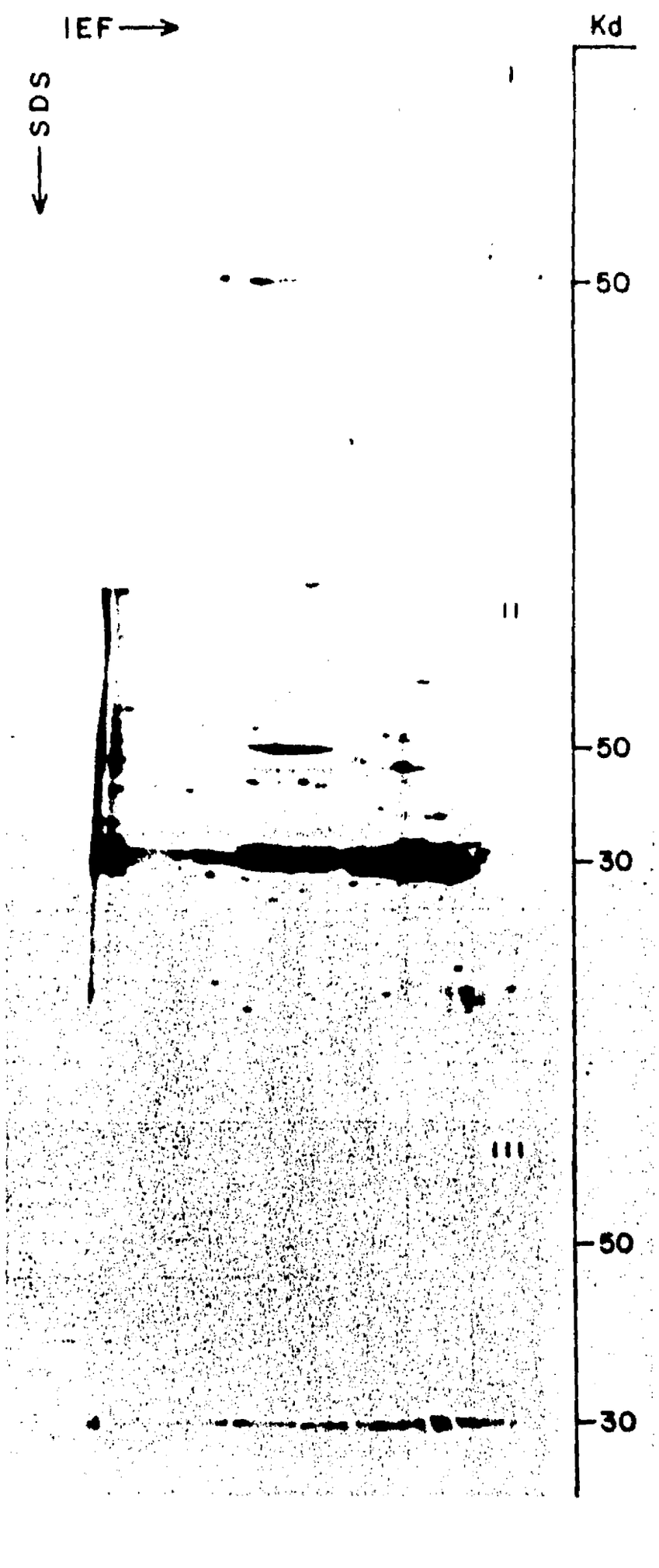

Fig. 4. Translation of hybrid-selected mRNA from the pNF-N30-1 cDNA clone. Polysomal RNA from 18-day nodules were processed for: (I), hybrid selection with pBR329; (II), *in vitro* translation without selection and (III), hybrid selection with digested cDNA of pNF-N30-1 clone. Released RNA was *in vitro* translated and ^{35}S-products were subjected to 2D-PAGE before fluorography.

'groups' of the 30 kDa nodulin cluster (see below). However, two major spots can be seen within the smear across the IEF gradient (Fig. 4-III). The isolation and characterization of the *P. vulgaris* nodule-specific uricase cDNA clone (pNF-UR07) has been described elsewhere [29].

Expression of the main nodule-specific transcripts during development of P. vulgaris *nodules*

The correlation between nodule age and the expression pattern of the 30 kDa nodulin (N-30) cluster was followed by 2D-PAGE of *in vitro* translation products (Fig. 5) and Northern blot analysis (Fig. 6). The presence of these N-30 gene products was first detected at day 13 after inoculation. By comparing with a group of acidic spots present in both root and nodule gels, it seems that the peculiar IEF pattern of N-30 cluster is constant throughout development. Nevertheless, the existence of two or three 'groups' or 'zones' in the streak can be assumed (see Fig. 5/Nodule 20-day-old), and the relative proportions of these 'groups' seem to change during late (16–28 days) nodule development. The temporal expression profiles of main nodulin mRNAs (uricase II, N-30 and leghemoglobin) indicated that all of these transcripts begin to accumulate coordinately from day 11 after inoculation (Fig. 6), and before the onset of nitrogen fixation [25]. pNF-N30-1 hybridized to several transcripts in the later (18–28 days) stages of nodule development, but no cross-hybridization was detected in these conditions with leaf RNA (lane H).

Expression of main nodulins in ineffective nodules of P. vulgaris

Nodule-specific soluble proteins and *in vitro*

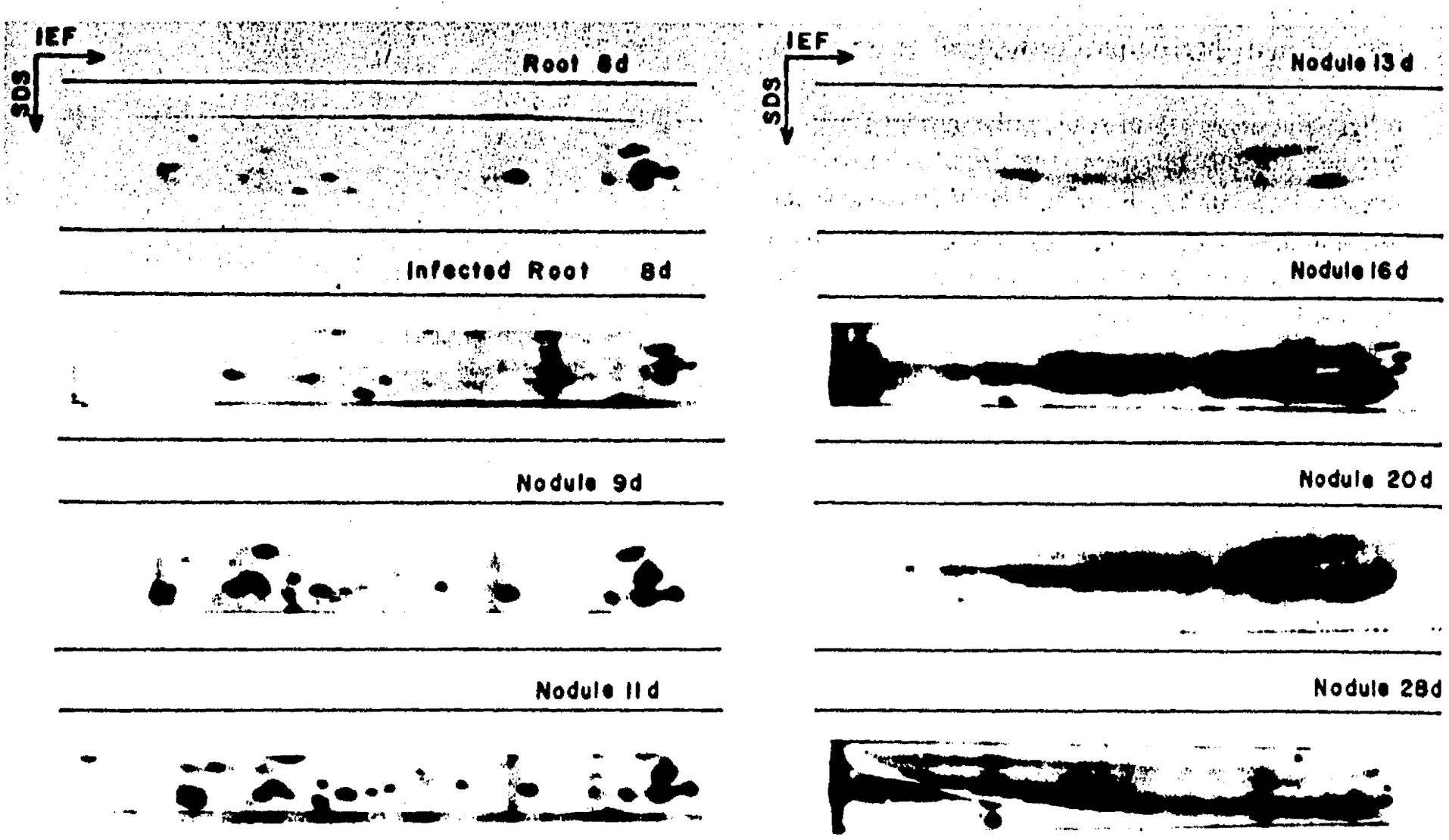

Fig. 5. Induction of nodulin-30 cluster gene products during nodule development. Two-dimensional analysis of *in vitro* translation products (about 5×10^5 cpm) of polysomal RNA from 8-day uninoculated and infected roots, and nodules of 9, 11, 13, 16, 20 and 28 days after inoculation. Only the 30 kDa zone of gels is shown.

products from wild type (Fix$^+$) and ineffective (Fix$^-$) nodules produced by *Rhizobium* sp. strains were compared. The relative expression level of the main soluble nodulins were decreased in ineffective nodules (Fig. 7A, lane 2). Few protein bands observed in the root extract (lane 3) were still induced in the nodules of CFN400. Conversely, the expression of a 150 kDa nodulin band was found to be increased in these nodules compared to effective ones. The SDS-PAGE analysis of products from the *in vitro* translation (Fig. 7B) revealed also that most nodulin transcripts were markedly decreased in ineffective nodules, particularly those corresponding to the N-30 cluster, to the uricase II and to a polypeptide in the 45–66 kDa range. It is interesting to note that none of the nodule-specific products in the *in vitro* translation of soybean RNA were equivalent to the induction level of N-30 (Panel B, lane 5), nor in the 2D-PAGE position (data not shown). Finally, the 2D pattern of *in vitro* translated products from ineffective bean nodules (Fig. 8) also showed the presence of true nodulins (large arrows), the persistence of root-specific products (small arrows) and also the lack of other nodulins (circles). This pattern reflected a disparity of some polypeptides from ineffective nodules relative to that from effective nodules of the same age (Fig. 3). The N-30 cluster was not clearly discernible in ineffective nodules, however, the zone covered by the streak resembles that of effective nodules during early (11–13 days) development (Fig. 5).

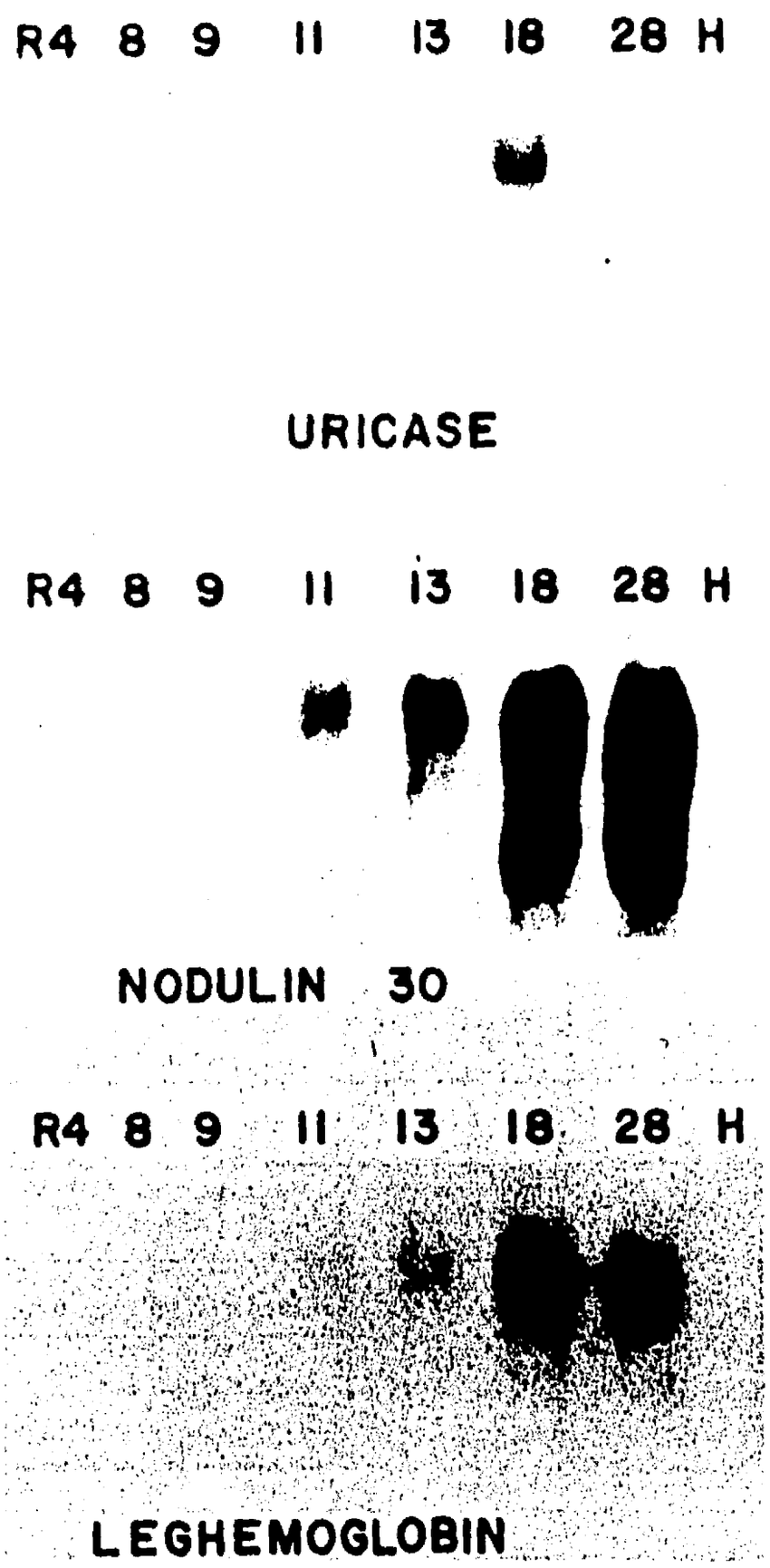

Fig. 6. Northern-blot analysis of major nodulin transcripts from *P. vulgaris* during the nodule development. Polysomal RNA from 4-day uninoculate roots (R4), developing nodules (8, 9, 11, 13, 18, 28 days after inoculation) and trifoliate leaves (H) were electrophoresed, blotted onto nitrocellulose and hybridized with the indicated probes (see Results).

Discussion

Specific proteins have been described to be induced during nodule organogenesis in legumes [34]. We report here the presence of about 20 nodulins of *Phaseolus vulgaris*. Some of these proteins play important metabolic roles in nodule functioning. Leghemoglobin (Lb) supports bacteroidal respiration and nitrogenase low-oxygen requirements [1]. The organ-specific glutamine synthetase [14, 25] is part of the ammonia assimilatory pathway and uricase II [29] contributes to the formation of ureides for export via the xylem. In *P. vulgaris* these are three of the most abundant nodule-specific proteins. The participation of other nodulins, as well as the mechanisms concerned with their regulation are to be understood. Moreover, nodulins could be involved in the biosynthesis of nitrogenous compounds, the catabolism of sucrose derivatives and as

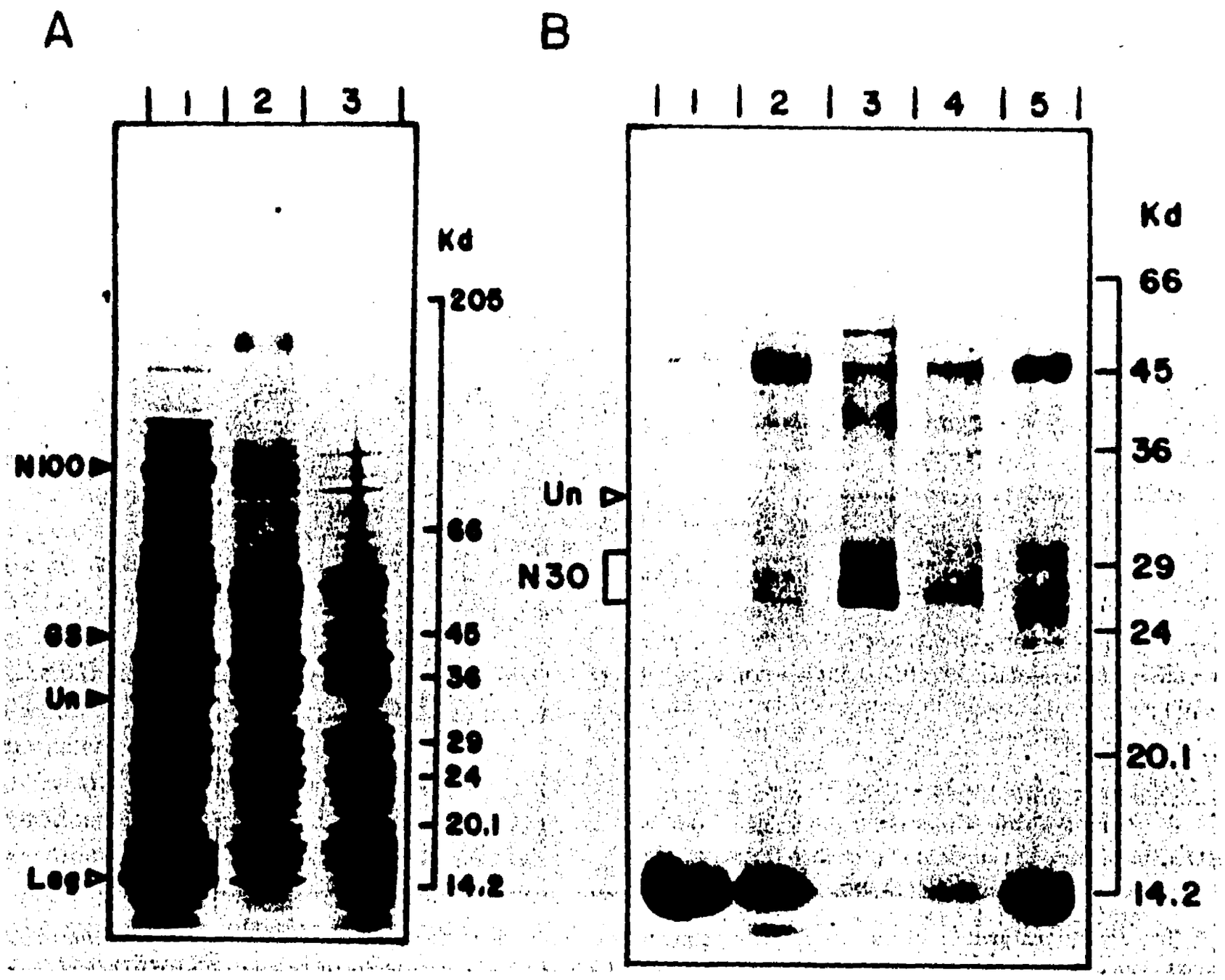

Fig. 7. Nodule-specific soluble proteins and *in vitro* translation products expressed in ineffective nodules of *P. vulgaris*. Panel A: 150 μg of soluble proteins from: (1), effective 18-day nodule; (2), 18-day ineffective (Fix⁻) nodule and (3) 8-day root fractions were subjected to 7.5–20% SDS-PAGE and stained with Coomassie blue. Positions of main nodulins are pointed as in Fig. 1. Panel B: Fluorogram of *in vitro* translation products obtained from a reticulocyte lysate programmed with: (1), no RNA added; 10 μg of polysomal mRNA from (2) 8-day roots; (3), effective 18-day nodules; (4), ineffective (Fix⁻) 18-day nodules and (5) 28-day soybean nodules. The 30 kDa cluster (N-30) and the uricase II (Un) products are indicated.

peribacteroidal membrane carriers in *P. vulgaris* and other legumes [33, 34]. In addition, during nodule development not only specific proteins are induced but also the repression of root-specific enzymes [25, 28] and the increase of enzymes which were previously present in roots, seem also to occur [27, 28].

A group of abundant nodulin transcripts that encode for proteins in the 30 kDa range were detected by *in vitro* translation from *P. vulgaris* nodule mRNA (Figs. 2 and 3). These nodulins exhibit a peculiar pattern in 2D-PAGE and neither RNAse nor SDS treatments improved its resolution. The 30 kDa (N-30) cluster might be encoded by a multigenic family as infered by two kinds of data: 1a) the 2D pattern of N-30 cluster suggests at least three different IEF zones in which various polypeptides could be resolved; 1b) the intensity of these zone differed during late nodule development (Fig. 5), although the possibility that this peculiar IEF pattern could be an artifact generated by the translation and/or the electrofocussing systems, still remains. 2a) The pNF-N30-1 cDNA clone hybrid-selected an mRNA which apparently encoded for only one or two of the membranes of this cluster (Fig. 4; 2b) this clone hybridized with multiple transcripts of nodule polysomal RNA (Fig. 6). By analogy, a nodulin gene family

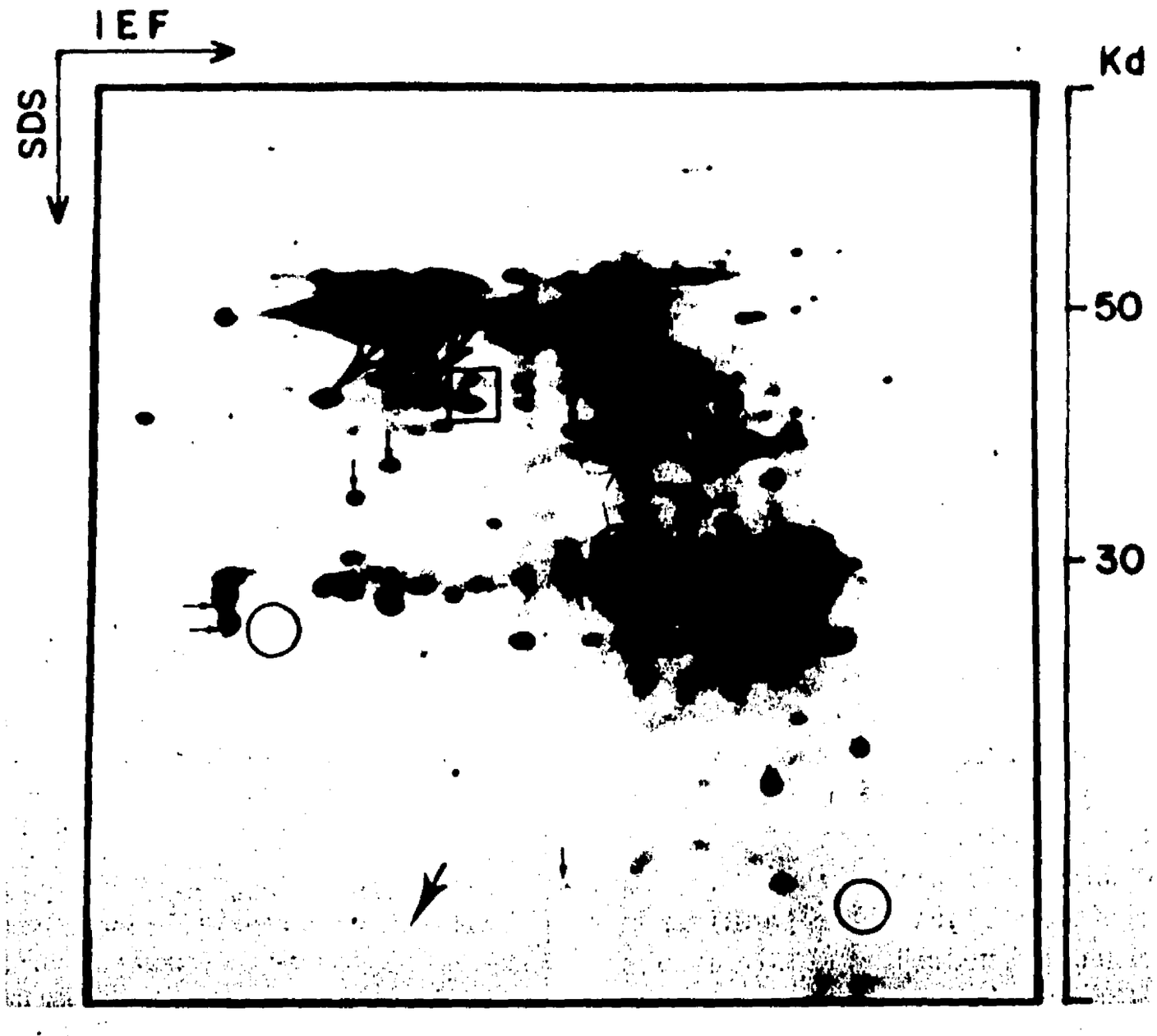

Fig. 8. Two-dimensional pattern of nodule-specific gene products expressed in ineffective nodules of *P. vulgaris*. *In vitro* translation products from polysomal RNA from 28-day nodules produced by CFN400 (Nod$^+$ Fix$^-$) were subjected to 2D-PAGE before fluorography. Symbols indicate: large arrows, nodule-specific products; circles, positions of lacking nodule-specific products; small arrows, root-specific products. GS-beta (enclosed in a square) was used as position marker.

which encodes for polypeptides of different MW has been described in soybean [30]. Whereas the N-30 cluster represents some of the most abundant nodulins, either as *in vitro* translation products (Fig. 3) or as transcripts (Fig. 6), any corresponding *in vivo* polypeptides with the same proportions have not been found, neither in the soluble nor in the peribacteroidal protein fractions (Figs. 1 and 7). This could be explained if these proteins had *in vivo* modifications and/or short half-life spans. Attempts are made to obtain antisera directed against various nodule fractions, in order to examine this idea. Expression of main *P. vulgaris* nodulins during development was approached by *in vitro* translation (Fig. 5) and Northern-blot procedures (Fig. 6). Results indicated that the N-30 cluster gene products appear from early organogenesis (day 13) and are detected uninterruptedly at least up to day 28 after inoculation (Fig. 5). Their corresponding transcripts were synchronously expressed along with leghemoglobin and uricase II mRNAs at day 11 of nodule development and before nitrogenase activity was detected (day 13). While N-30 transcripts accumulated significantly, Lb and uricase II seem to decline in 28-day nodules. The appearance of the N-30 cluster as *in vitro* translation products was first detected at day 13, but at day 11 by hybridization; this was probably due to dissimilar sensitivity thresholds between these procedures.

In order to ascertain the relation between nodule effectiveness and the expression of main nodulins, the relative level of the *in vivo* (Fig. 7A) and *in vitro*

(Figs. 7B and 8) synthesized proteins in ineffective (Fix$^-$) nodules was compared. Reduced amounts of main nodulins, as well as the persistence of some root-induced and specific proteins in nodules of CFN400 suggested that the developmental sequence of nodule gene expression was somehow arrested in a stage similar to that of early days of nodulation and also that the organogenesis was impaired. These statements could be stressed by the fact that nodules formed by CFN400 not only lacked nitrogenase activity, but contained low RNA levels and were constituted by very few infected cells (unpublished observations). There was enough homology between some nodulins of soybean to allow the isolation of corresponding cDNA clones for leghemoglobin and uricase II [29]. However, there is no cross-hybridization of soybean nodulin-23 and -24 cDNA clones with *P. vulgaris* nodule RNA [7]. It is therefore likely that genes encoding for the N-30 cluster and their developmental induction may be a particular character of *P. vulgaris* nodules.

In view of these observations we postulate that gene products of the N-30 gene cluster participate in the development of infected cells of effective nodules and that they could be induced in close association with the intracellular infection by *Rhizobium*. We also conclude that the induction of some nodulins is not coordinated and is regulated at different levels and/or by diverse factors. The nature of some of these mechanisms are becoming clearer in other symbiotical systems [9, 15, 34]. Work is in progress to obtain the primary structure of full-length pNF-N-30 cDNA clones and to localize these products in subcellular nodule compartments.

Acknowledgements

We thank L. Segovia, R. Palacios and A. Covarrubias for the critical reading of this manuscript. To W. J. Swain for helping us to construct the cDNA library and D. P. S. Verma for providing soybean leghemoglobin and pNod24cDNA clones; to A. Garciarrubio for helpful tips in Northern-blotting and J. L. Zitlalpopoca, A. Pichardo, R. Basurto and D. Cuéllar for technical and secretarial help.

References

1. Appelby CA: Leghemoglobin and *Rhizobium* respiration Ann Rev Plant Physiol 35: 443–478, 1984.
2. Atkins CA: Ureide metabolism and the significance of ureides in legumes. In: Subba-Rao NS (ed) Advances in Agricultural Microbiology. Oxford and IBH, New Dehli, 1982, pp 25–51.
3. Bergmann H, Preddie E, Verma DPS: Nodulin-35: a subunit of specific uricase (uricase II) induced and localized in the uninfected cells of soybean nodules. EMBO J 2: 2333–2339, 1983.
4. Brewin NJ, Robertson JG, Wood EA, Wells B, Larkins AP, Galfre G, Butcher GW: Monoclonal antibodies to antigens in the peribacteroid membrane from *Rhizobium*-induced root nodules of pea cross-react with plasma membranes and Golgi bodies. EMBO J 4: 605–611, 1985.
5. Christofferson RE, Laties GG: Ethylene regulation of gene expression in carrots. Proc Natl Acad Sci USA 79: 4060–4063, 1982.
6. Fuller F, Künstner PW, Nguyen T, Verma DPS: Soybean nodulin genes: analysis of cDNA clones reveals several major tissue-specific sequences in nitrogen-fixing root nodules. Proc Natl Acad Sci USA 80: 2594–2598, 1983.
7. Fuller F, Verma DPS: Appearance and accumulation of nodulin mRNAs and their relationship to the effectiveness of root nodules. Plant Mol Biol 3: 21–28, 1984.
8. Govers F, Gloudemans T, Moerman M, van Kammen A, Bisseling T: Expression of plant genes during the development of pea root nodules. EMBO J 4: 861–867, 1985.
9. Govers F, Moerman M, Downie JA, Hooykaas P, Franssen HJ, Louwerse J, van Kammen A, Bisseling T: *Rhizobium nod* genes are involved in inducing an early nodulin gene. Nature 323: 564–567, 1986.
10. Graham P: Characteristics of selected strains of *Rhizobium phaseoli* from CIAT collection. In: CIAT *Rhizobium phaseoli* collection. Cali, 1980, pp 29–30.
11. Jackson AO, Larkins BA: Influence of ionic strength, pH, and chelation of divalent metals on isolation of polyribosomes from tobacco leaves. Plant Physiol 57: 5–10, 1976.
12. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. Nature 227: 680–689, 1970.
13. Lang-Unnasch N, Ausubel FM: Nodule-specific polypeptides from effective alfalfa root nodules and from ineffective nodules lacking nitrogenase. Plant Physiol 77: 833–839, 1985.
14. Lare M, Porta H, Padilla J, Folch J, Sánchez F: Heterogeneity of glutamine synthetase polypeptides in *Phaseolus vulgaris* L. Plant Physiol 76: 1019–1023, 1984.
15. Larsen K, Jochimsen BU: Expression of nodule-specific uricase in soybean callus tissue is regulated by oxygen. EMBO J 5: 15–19, 1986.
16. Laskey RA, Mills AD: Quantitative film detection of ^3H and ^{14}C in polyacrylamide gels by fluorography. Eur J Biochem 56: 335–341, 1975.

17. Lee JS, Verma DPS: Sructure and chromosomal arrangement of leghemoglobin genes in kidney bean suggest divergence in soybean leghemoglobin gene *loci* following tetraploidization. EMBO J 3: 2745–2752, 1984.
18. Legocki RP, Verma DPS: Identification of 'nodule-specific' host proteins (nodulins) involved in the development of *Rhizobium*-legume symbiosis. Cell 20: 152–163, 1980.
19. Lowry OH, Rosebrough NJ, Farr A, Randall RJ: Protein measurements with Folin phenol reagent. J Biol Chem 183: 265–275, 1951.
20. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. CSH Laboratory, Cold Spring Harbor, New York, 1982.
21. Martinez E, Flores M, Brom S, Romero D, Dávilla G, Palacios R: *Rhizobium phaseoli:* a molecular genetics view. Plant & Soil, 1987 (in press).
22. Mellor RB, Christensen TMIE, Werner D: Choline kinase II is present only in nodules that synthesize stable peribacteroid membranes. Proc Natl Acad Sci USA 83: 654–663, 1986.
23. Newcomb W: Nodule morphogenesis and differentiation. In: Giles K, Atherly AG (eds) International Review of Cytology, Supplement 13. Academic Press, New York, 1981, pp. 247–298.
24. O'Farrell PH: High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250: 4007–4021, 1975.
25. Padilla JE, Campos F, Conde V, Lara M, Sánchez F: Nodule-specific glutamine synthetase in expressed before the onset of nitrogen fixation in *Phaseolus vulgaris* L. Plant Mol Biol 9: 65–74, 1987.
26. Pelhalm HRB, Jackson RJ: An efficient mRNA dependent translation system from reticulocyte lysates. Eur J Biochem 67: 247–256, 1976.
27. Reynolds PHS, Boland MJ, Belvins DG, Schubert KR, Randall DD: Enzymes of amide and ureide biogenesis in developing soybean nodules. Plant Physiol 69: 1334–1338, 1982.
28. Robertson JG, Farnden KJF: Ultrastructure and metabolism of the developing legume root nodule. In: Miflin BJ (ed) The Biochemistry of Plants, Vol 5. Academic Press, New York, 1980, pp 65–113.
29. Sánchez F, Campos F, Padilla J, Bonneville JM, Enríquez C, Caput D: Purification, cDNA cloning and developmental expression of the nodule-specific uricase from *Phaseolus vulgaris* L. Plant Physiol 84, 1987 (in press).
30. Sengupta-Gopalan C, Pitas JW, Thompson DV, Hoffman LM: Expression of host gene during root-nodule development in soybeans. Mol Gen Genet 203: 410–420, 1986.
31. Thomas PS: Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc Natl Acad Sci USA 77: 5201–5205, 1980.
32. Thummler F, Verma DPS: Nodulin-100 gene of soybean codes for nodule-specific sucrose synthetase. In: Third International Symposium on the Molecular Genetics of Plant-Microbe Interactions. Montreal, 1986, p 182.
33. Verma DPS, Long S: Molecular biology of *Rhizobium*-legume symbiosis. In: Jeon K (ed) International Review of Cytology, Supplement 14. Academic Press, New York, 1983, pp 211–245.
31. Verma DPS, Fortin MG, Stanley J, Mauro VP, Purohit S, Morrison N: Nodulins and nodulin genes of *Glycine max:* a perspective. Plant Mol Biol 7: 51–61, 1986.
35. Vincent JM: Factors controlling the *Rhizobium*-legume symbiosis. In: Newton WE, Orme-Johnson H (eds) Nitrogen Fixation. University Park Press, Baltimore, 1980, pp 103–136.

II. RESULTADOS.

El desarrollo de este trabajo puede dividirse en dos etapas: en la primera se detectó la existencia del grupo de la Npv30 de nódulos de frijol, los resultados obtenidos se han reportado en el articulo "**Expression of nodule-specific genes in *Phaseolus vulgaris* L.**" y esta contenido en la sección de Antecedentes. En la segunda etapa se llevo a cabo una caracterización mas exhaustiva de este grupo de nodulinas. En esta sección se muestran los resultados obtenidos de la caracterización de la familia de Npv30, los cuales se reportarón en:

Campos F, Carsolio C, Kuin H, Bisseling T, Rocha-Sosa M, Sánchez F. (1995). **Characterization and gene expression of nodulin Npv30 from Phaseolus vulgaris L.** Aceptado para su publicación en Plant Physiology.

Running title: **Npv30 expression in common bean nodules**.

Corresponding author:

Federico Sánchez

Departamento de Biología Molecular de Plantas

Instituto de Biotecnología

Universidad Nacional Autónoma de México

Apdo. Postal 510-3

Cuernavaca, Morelos 62271

México

Phone: (52) (73) 11-40-00

Fax: (52) (73) 17-23-88.

PLANT PHYSIOLOGY

1

Published by the American Society of Plant Physiologists

Editor-in-Chief, Maarten J. Chrispeels

Associate Editors:
Gloria M. Coruzzi
Malcolm C. Drew
Thomas J. Guilfoyle
Sharon R. Long
William L. Ogren
Charles F. Yocum

ANN M. HIRSCH
Editor

3 March 1995

Dr. Federico Sánchez
Departamento de Biología Molecular de Plantas
Instituto de Biotecnología
Universidad Nacional Autónoma de México
Apdo. Postal 510-3
Cuernavaca, Morelos 62271, México

Dear Federico:

Your manuscript, P95-0003, has been reviewed by two experts and I have served as third reviewer. The reviews and annotated manuscripts are enclosed with this letter; I have kept one copy of the manuscript for my records.

I have provisionally accepted this manuscript pending revision. Both reviewers found that the manuscript was much improved over its previous incarnation. You will note that I have heavily edited one copy of your manuscript. Reviewer #1 has also included numerous editorial marks and corrections. These should help you fix up the manuscript so that you can address some of Reviewer #2's concerns about the grammatical errors and spelling mistakes. I disagree with Reviewer #1 about changing the title because I agree strongly with Reviewer #2 that you have not demonstrated whether or not the weak signal could be attributed to the low titer of the antibody or due to the fact that the antibody was raised to a fusion protein. Have you done the controls suggested by Reviewer #2? If so, please mention this in a cover letter to me so I can allay my concerns about the antibody. Reviewer #1 also points out that you have not discussed any of the controls for the antibody experiments. I am not certain why the Reviewer thought you were trying to immunolocalize this nodulin on sections, so ignore that statement, but I do think it is important that you let me know that the proper controls were performed. It might also be a good idea to include a statement in the Results section.

On p. 15, I think I know what you are trying to say about the protein-protein interactions (i.e., Npv30-1 and cell wall proteins), but I am not sure. I have reworded this sentence a bit so please check that I have kept the meaning that you wanted.

On p. 17, Reviewer #1 wants you to address the "errors" of *in situ* hybridization but I don't think this is necessary in your paper. I see the Reviewer's point, but the task is much more global than your paper. Perhaps someone should address the issue of transcript abundance and protein levels in a workshop, in a Proceedings article, or in a review, but not in a research paper.

Department of Biology, 405 Hilgard Avenue, University of California, Los Angeles, Los Angeles, CA 90024-1606 USA
Telephone 310-206-8673, Fax 310-206-3987

95

5413

To be honest, I hope that Fig. 8A and B look better in the original than they do in the copy. It is very difficult to see the cell outlines in Fig. 8A. I presume the blue "spots" within the infected cells are the nuclei, n'est ce pas? *

I have not had the time to check to see that all the references cited in the manuscript are actually present in the Literature Cited. Please do this for your revised copy. Also, please update any articles listed as "in press".

You have 60 days to return a revised manuscript and one copy to me. Also, please return the copy with my and Reviewer #1's comments so I can see that you have made the changes I have recommended. If you disagree with any of these recommendations, please state so in your cover letter to me. If you have any questions about my letter, the manuscript, etc., please do not hesitate to contact me at any one of the numbers given below.

Muchas gracias, Federico, por mandar su manuscrito á Plant Physiology. No es facile escribir en otro idioma y ustedes lo han hecho muy bien. Yo espero que yo escriba español tan bien!

Very truly yours,

Ann M. Hirsch
Monitoring Editor and
Professor of Biology

Enclosures

* My French is worse than my español.

Characterization and Gene Expression of Nodulin Npv30 from *Phaseolus vulgaris* L.

Francisco Campos, Carolina Carsolio, Hilda Kuin[1], Ton Bisseling[1], Mario Rocha-Sosa and Federico Sánchez*.

Date received: Date accepted:

Departamento de Biología Molecular de Plantas
Instituto de Biotecnología
Universidad Nacional Autónoma de México
Apdo. Postal 510-3
Cuernavaca, Morelos 62271
México

[1]Department of Molecular Biology
Agricultural University
De Dreijen 11
6703 BC Wageningen
The Netherlands

ABSTRACT

We have previously reported that transcripts for a 30 kDa nodulin (Npv30) are very abundant in the nodule. This paper describes the isolation and characterization of Npv30 cDNA and genomic clones. Npv30 has the following characteristic features: a) a putative signal sequence at the deduced amino terminal region; b) a proline-rich stretch at the carboxy-terminal end; and c) a characteristic domain of four cysteines resembling metal-binding sites. In *Phaseolus vulgaris* L. Npv30 is encoded by a small gene family which shares discrete sequence homologies with another small gene family in soybean. An antibody against a β-galactosidase-Npv30 fusion protein detected two proteins of 28 and 30 kDa. Although Npv30 transcripts are very abundant, they encode proteins that were hardly detected in nodule fractions, suggesting that these proteins have a short half-life and/or the mRNAs are strongly regulated at the translational level. By *in situ* hybridizations experiments, Npv30 transcripts were detected in the infected cells of the nodule.

INTRODUCTION

The nitrogen-fixing nodule is a specialized organ induced on the roots or stems of leguminous plants by Gram-negative soil bacteria of the genera *Rhizobium*, *Bradyrhizobium* or *Azorhizobium*. During the development and function of this organ a set of both rhizobial and plant host genes are specifically expressed (Hirsch, 1992; Fisher and Long, 1992; Sánchez *et al.*, 1991).

Several nodule-specific plant genes which encode proteins called nodulins are activated during rhizobial infection, nodule morphogenesis, and the establishment of the appropriate environment for nitrogen-fixation (Legocki and Verma, 1980; van Kammen, 1984). According to their timing of appearance during nodule development, these genes have been classified into early and late nodulin genes (Govers *et al.*, 1987; Nap and Bisseling, 1990). Genes encoding early nodulins are induced during rhizobial infection, growth of the infection thread and stimulation of cortical cell divisions (Franssen *et al.*, 1992; Kouchi and Hata, 1993). The expression of some early nodulin genes, such as ENOD12 can be induced by rhizobial lipooligosaccharides known as nodulation (Nod) factors (Journet *et al.*, 1994; Pichon *et al.*, 1992). However, others as ENOD2, cannot be induced by the Nod factors and they seem to be under hormonal regulation (Dehio and Bruijn, 1993; Cooper and Long, 1994).

Late nodulin genes are activated in the developing and mature nodule concomitant with the onset of nitrogen fixation activity. There are reports of late nodulin genes that appear to be activated later in nodule development. A nodule-specific protease inhibitor is expressed in senescent nodules of winged bean (Manen *et al.*, 1991). Also, in pea a late nodulin gene, PsNOD6, is activated in the infected cells following the induction of the leghemoglobin genes (Kardailsky *et al.*, 1993). In soybean and common bean, several late nodulin genes have been studied and the function of some of their products have been described; for example: leghemoglobin (Appleby, 1984; Lee and Verma, 1984; Campos *et al.*, 1987), the gamma subunit of glutamine synthetase from *P. vulgaris* (Lara *et al.*, 1983; Padilla *et al.*, 1987), uricase II (Bergmann *et al.*, 1983, Sánchez *et al.*, 1987), and sucrose synthase (Thummler and Verma, 1987) among others.

A highly expressed late nodulin gene family of unknown function, the nodulin-A family, has been found in soybean (Fuller *et al.*, 1983; Jacobs *et al.*, 1987). The nodulin-A family comprises at least six members: Ngm-20 (Sandal *et al.*, 1987), Ngm-23 (Mauro *et al.*, 1985), Ngm-26b (Jacobs *et al.*, 1987), Ngm-22/Ngm-27 (Jacobs *et al.*, 1987; Sandal *et al.*, 1987), Ngm-44 (Sengupta-Gopalan *et al*, 1986) and 15-9-A (Gottlob-McHugh and Johnson, 1991). The members of this gene family show two common characteristics: (a) two domains that are arranged in paired Cys-X7-Cys motifs, resembling zinc-

finger sequences ; and (b) a putative signal peptide (Sandal *et al.*, 1987).

In *P. vulgaris*, a group of abundant late nodule specific transcripts that encode proteins in the 30 kDa range (Npv30) were identified by *in vitro* translation of hybrid-released mRNA (Campos *et al.*, 1987). Npv30 exhibits a peculiar electrofocusing pattern which appeared as a streak across the IEF gradient (pH 3-10). These results suggested that the 30 kDa nodulin group could be derived from multiple related transcripts encoding polypeptides of similar molecular weight (Campos *et al.*, 1987).

Here we show that Npv30 is indeed encoded by a gene family similar to the well-characterized nodulin gene family-A of soybean, described by Jacobs *et al.*, (1987). We were able to detect nodulin Npv30 with antisera against β-galactosidase-Npv30 in soluble nodule fractions and by using *in situ* hybridization, we concluded that the Npv30 transcripts were expressed in the infected cells of the nodule.

MATERIAL AND METHODS.

Plant material.

Phaseolus vulgaris L. seeds cv. Negro Jamapa, were obtained from PRONASE, Mexico. Seeds were germinated for 2 d, transferred to pots, inoculated with *Rhizobium tropici* strain CIAT 899 (Martínez-Romero *et al.*, 1991) and grown in the greenhouse as previously described by Lara *et al.* (1984). Nodules were harvested 21 d after inoculation. Uninfected bean roots were obtained and collected 6 d after sowing. Both

tissues were harvested and frozen immediately in liquid nitrogen and stored at -70°C.

Isolation of plant DNA and Southern blot analysis.

Genomic bean DNA was isolated from leaves by the CTAB method described by Doyle and Doyle (1990). Approximately 5 µg DNA were digested, subjected to electrophoresis on 0.8% agarose gel and transferred to Nylon filters. The blot was hybridized at 42° C to random-primed probes in 50% formamide according to Sambrook *et al.*(1989). The blots were washed three times for 15 min each in 0.1 X SSC, 1% SDS at 65° C.

Northern dot blot analysis.

Total RNA from 8-d-old roots and 21-d-old nodules was prepared by the "hot-phenol" method, following the procedure of de Vries *et al.* (1991). Serial dilutions of root and nodule RNA were dotted on to Nylon filters, and hybridized to leghemoglobin and Npv30 cDNAs. Hybridization and washes were done at a high stringency as previously described by Sambrook *et al.* (1989). After hybridization of northern dot blots, the filters were autoradiographed and the intensity of spots was measured with a Bio Image System (Millipore Co., Ann Arbor, MI) using Visage software.

DNA sequencing.

Appropriate DNA restriction fragments of cDNA and genomic Npv30 clones were subcloned into the pKS^+ and pSK^+ vectors (Stratagene, La Jolla, CA) and sequenced using Sequenase kit (USB Corp., Cleveland, OH) following the instructions of the manufacturer. Nucleotide sequences have been submitted to EMBL

data bank under the accession numbers Z30345 for the cDNA pN311, and Z30347 for the genomic clone *npv*30-1. Sequence data were analyzed using the software package of the University of Wisconsin Computer Group.

Production of Npv30 fused to β-galactosidase protein in *E. coli*.

A gene fusion between the 3' terminus of *lacZ* and part of the Npv30 cDNA clone was made using pUR289 as expression vector (Rhüter and Müller-Hill, 1983; Sambrook *et al.*, 1989). A 594 bp *RsaI- RsaI* fragment of the pNF311 clone coding for the last 159 amino acids of the carboxy terminal-end (see Fig 2) was subcloned into the *Sma*I site of pKS$^+$ (pNF308 clone). A *BamHI-HindIII* fragment of this clone was introduced into the pUR289 plasmid to produce the pSYC30 plasmid. *E. coli* JM109 was transformed with pSYC30 and a fusion protein was induced as recommended by Sambrook *et al.*(1989). The β-galactosidase-Npv30 fusion protein was purified by affinity chromatography using a p-aminobenzyl 1-thio beta-D-galactopyranoside agarose column (Das, 1990). Antibodies against this fusion protein were produced in a rabbit by standard methods.

Protein extraction.

Total proteins from 8-d-old roots and 21-d-old nodules were extracted by grinding 1 g of frozen tissue and homogenized in 2 mL of 100 mM Tris-HCl, pH 6.8, 5% (w/v) SDS, 5 mM 2-mercaptoethanol. The homogenate was boiled for 5 min and centrifuged at 1 x 10^4 g for 5 min. The supernatant is referred to as "total fraction". Freshly extracted roots and

nodules were used to prepare soluble protein fractions. One g of tissue was gently ground in 100 mM Tris-HCl pH 8.0, 0.5 M sucrose and 5% (w/v) polyvinyl polypyrrolidone. The homogenate was centrifuged at 1 x 10^4 g for 15 min and the supernatant was considered to be a "soluble fraction".

Western-blot detection.

Proteins were separated by 15% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) as described by Laemmli (1970). Protein gel blotting on Immobilon (Millipore Co., Bedford, MA) was carried out as described by Towbin *et al.* (1979). The immunodetection was performed using secondary antibodies conjugated with alkaline phosphatase (Blake *et al.*, 1984). The preimmune serum did not recognize neither Npv30 in a nodule extract nor pure beta galactosidase. The antiserum recognises, in the same blot, Npv30 in a nodule extract and beta galactosidase as one of the molecular weight markes at a 1: 5000 dilution.

***In situ* hybridization.**

Twenty-one-d-old nodules were fixed, dehydrated, and embedded into paraffin according to procedures described by van de Wiel *et al.*, (1990). Nodule sections were hybridized with [^{35}S]UTP-labeled antisense or sense RNA probes (100- 1500 Ci/ mmol) made from the pN308 plasmid according to the conditions reported by Scheres *et al.* (1990). Sections were stained with 0.025% toluidine blue after 1 week exposure at 4° C and photographed with a microscope equipped with dark-field and epipolarization optics.

RESULTS.

Isolation of the Npv30 cDNA and gene.

A nodule cDNA library of *P. vulgaris* L. Negro Jamapa made in pBR322 (Campos *et al.*, 1987) was screened using the 700 bp *PstI* insert of pNF-30-1 (Campos *et al.*, 1987). Attempts to isolate a full-length Npv30 cDNA clone in this library were unsuccessful. A partial 1-kb cDNA clone from Npv30 (pN311) was isolated, sequenced and used for further analysis. The genomic λ123 clone was isolated from a *P. vulgaris* L. Saxa library (Clontech, Palo Alto CA) as previously described (Carsolio *et al.*, 1994). The restriction map of the genomic λ123 clone, showing that *npv*30-1 gene is contained within a 7.2 kb *SalI* genomic fragment, is presented in Fig.1. The 2.1 kb *SalI*-*Hind*III fragment from 1123 that hybridized to a 5' specific probe from pN311 was subcloned in the vector pKS$^+$, giving rise to the plasmid pGN338 (Fig. 1). This 2.1 kb *SalI*- *Hind*III fragment was completely sequenced. The analysis of the sequence revealed the presence of a putative promoter, an exon (exon 1) containing most of the coding region, and part of an intron, as compared with the sequence of pN311 clone. The sequence analysis of the control region of this gene has been reported by Carsolio *et al.*(1994). To complete the sequence of the intron and the other exon (exon 2), we constructed the plasmid pGN339 (Fig. 1), by introducing the 7.2 kb *SalI* fragment into pKS$^+$. The sequencing reactions were primed with specific oligonucleotides.

Sequence Analysis.

The nucleotide sequences of the partial length cDNA (pN311) and genomic subclones are shown in Fig. 2. The nucleotide sequence of Npv30 cDNA includes a 663-bp open reading frame, a 175-bp 3' untranslated region containing a putative polyadenylation signal and a poly(A^+) tail (Fig. 2).

The deduced protein sequence of the genomic clone (refered to as Npv30-1) corresponds to a polypeptide of 220 amino acid residues and a molecular mass of 23.6 kDa. Hydropathy analysis determined by the method of Kite and Doolittle (1982) (Fig. 3) predicted that the amino-terminal of Npv30-1 is highly hydrophobic and sequence analysis of this region suggested a potential cleavage site for the putative signal peptide (Fig. 2) (von Heijne, 1986).

The deduced amino acid sequence analysis revealed high homology to discrete regions of the previously mentioned late nodulin-A gene family from soybean(see Fig. 4) (Jacobs *et al*, 1987; Sandal *et al.*, 1987). Both the bean gene and the soybean gene family exhibited conserved regions that could encode for putative signal peptides and the two domains arranged in paired Cys-X_7-Cys motifs, resembling zinc-finger sequences or metal-binding domains (Sandal *et al.*, 1987).The highest score of identity and similarity of the deduced amino acid sequence from the genomic clone was to 53.71% identity and 71.42% similarity to Ngm-20 and 47% identity and 65.51% similarity to Ngm-22/Ngm-27.

The comparison of the inferred amino acid sequences between both, cDNA and genomic clones (Fig. 2) showed that the cDNA encoded for a larger protein with 19 additional amino acids and more proline residues at the carboxy-terminal. Other base changes and mismatches between both clones can also be observed in Figure 2.

Genomic Blot Analysis.

To estimate the number of Npv30 genes in common bean, Southern analysis was performed. The 1-kb *PstI* insert from pN311 was used to probe blots of *P. vulgaris* genomic DNA digested with several enzymes (Fig. 5). At high stringency (0.1 x SSC, 65° C), the probe hybridized to at least three discrete DNA bands. These data together with the alignment of the genomic and the Npv30 cDNA sequences suggest that Npv30 is encoded by a small gene family.

Northern Dot Blot Analysis.

To quantify Npv30 mRNA in the nodule, dot blots of total RNA from 21-d-old nodules of common bean were probed at high stringency with both Npv30 and leghemoglobin cDNAs (Campos *et al.*, 1987). The lenght of both probes was of similar size (about 600 pb). Densitometric evaluation of the autoradiography showed that Npv30 transcripts are about 3-fold more abundant than leghemoglobin transcripts (Fig. 6).

Detection of Npv30 in the Nodule.

A *lac Z*- Npv 30 gene fusion was constructed by subcloning an internal pN311 *RsaI*- *RsaI* fragment in pUR289 (see Material and

Methods). Such a fusion protein lacked the 81 amino acid residues from the amino-terminal of Npv30. An antiserum was raised against β-galactosidase-Npv30 fusion protein. The antiserum reacted with total and soluble nodule fractions. The antibody detected two bands of approximately 30 and 28 kDa. .No signal was obtained with a leaf (data not shown) or root extracts(Fig. 7)

Localization of the Npv30 transcripts in the Nodule.

The localization of the Npv30 transcripts in nodules was analyzed by *in situ* hybridization. Sections from bean nodules, of 21-d-old plants, were hybridized with ^{35}S-UTP labeled RNA probes transcribed from the pN311 insert (sense and antisense). In Fig. 8, only the hybridization with the Npv30 antisense probe is presented. Npv30 transcripts accumulated to high levels in the infected cells, whereas no hybridization was detected in the uninfected cells of the nodule central zone (Figure 8 B,D)

Discussion.

We have previously described a group of abundant nodule-specific transcripts in *P. vulgaris* which *in vitro* translate proteins in the 30 kDa range (Npv30). We were also able to isolate a cDNA clone which hybrid-selected an mRNA of the Npv30 group (Campos *et al.*, 1987).

In this study, we have continued the characterization of Npv30. Although the cDNA and a genomic clones were isolated from different cultivars, the nucleotide sequences indicated that the Npv30 cDNA was derived from a different gene to the

*npv*30-1 genomic clone. Nevertheless, the comparison of these clones showed that they were highly homologous (95% identical). These data, together with additional results, such as: a) the electrofocusing pattern of *in vitro* translated products from hybrid-selected mRNA by pNF-30-1 clone (Campos *et al.*, 1987); b) Southern blot analysis (Fig.5); and c) the partial sequence of another cDNA clone different from N311 (data not shown), suggest that Npv30 is encoded by a gene family.

Analysis of the deduced amino acid sequence revealed that the Npv30 had homology with a highly transcribed nodulin-A family from soybean (Richter *et al.*, 1991; Gottlob-Mchugh and Johnson, 1991). Like all members of this soybean family, Npv30 had two regions which contained Cys-X7-Cys motifs. Cysteine residues in Npv30 may be involved in forming inter- or intramolecular disulfide bonds. Although it has been reported that proteins with regularly spaced cysteines may be metal-binding proteins (Berg, 1990), searches of protein databases did not show homology between Npv30 and any metal-binding protein so far reported. The conservation of these domains in both common bean and soybean suggests a similar function for these nodulins.

Another intriguing feature is the presence of a proline-rich carboxy-terminal end in Npv30 and some members of the soybean family, such as Ngm-20 and 15-9-A (Sandal *et al.*, 1987; Gottlob-Mchugh and Johnson, 1991).It is interesting to note that pN311 cDNA encodes for 8 additional proline residues at the carboxy-terminal as compared to the *npv30*-1 clone. These

proline residues could be involved in creating an adequate environment for protein-protein interaction (Yu *et al.*, 1994). Also, a loose interaction of these proteins with cell wall proteins through the prolines (hydroxyprolines?) cannot be discarded (Showalter, 1993).

The observation that *npv30*-1 like the Ngm-20, Ngm-22 (Sandal *et al.*, 1987), Ngm-23 genes (Mauro *et al.*, 1985), are interrupted by an intron at an equivalent position, suggests that the bean *npv30*-1 and soybean nodulin-A family have evolved from a common ancestor.

The fact that the *Npv30*-1 gene, as well as all members of the nodulin-A soybean gene family, encode for proteins with a putative signal peptide, suggests that these proteins may be associated to membranes or secreted from the cell, possibly to the peribacteroid space. Although the Npv30 proteins were detected in the soluble fraction (Fig. 7), different members of the nodulin-A gene family have different cellular localizations: some have been detected in the cytosol (Ngm-44; Ngm-22), whereas others can be located both in the cytosol and in the peribacteroid membrane (Ngm-26). Some others have been associated to the peribacteroid membrane (Ngm-20, Ngm-23) (Jacobs *et al.*, 1987; Richter *et al* 1991). This indicates that, it is necessary to carry out a more detailed and careful analysis of the subcellular localization of the Npv30 protein.

Although pN311 cDNA is a partial clone, we might predict that both the genomic *npv30* -1 and pN311 encode for proteins of similar size at the amino terminal end (see Fig. 2).

Nevertheless, pN311 should encode for a protein of about 2 kDa larger than the Npv*30*-1 protein, because the former has 19 additional amino acids (Fig. 2).

When the antibody directed to the β-galactosidase-Npv30 fusion protein from *E. coli* was used against total and soluble nodule bean fractions, two proteins of 30 and 28 kDa were detected by western blot analysis. These proteins may be the unprocessed and the matured protein after the removal of the signal peptide, or alternatively, the antibody detects pN311 and *Npv30*-1 products and/or other members of the Npv30 family.

There is discrepancy between molecular weight of the derived amino acid sequence and the molecular weight of the immunodetected proteins in the nodule. This phenomenon has already been observed for Ngm-22, Ngm-23 (Jacobs *et al.*, 1987; Richter *et al.*, 1991), and Ngm-24 (Katinakis and Verma, 1985)

In spite of that Npv30 mRNAs are very abundant, it was difficult to immunodetect these proteins in total or soluble nodule fractions. It is worth to mention that we previously reported (Campos *et al.*, 1987 and unpublished results) we were unable to detect abundant proteins between 20-30 kD range in soluble, peribacteroidal or total nodule fractions, as compared to leghemoglobin (20% of nodule protein), uricase II (2% of nodule protein) and glutamine synthetase (2% of nodule protein) when analyzed in one or two dimentional polyacrylamide gels silver-stained or with Coomassie. This suggests that the Npv30 proteins could have a short-half life and/or the mRNAs are strongly regulated at the translational level.

From the *in situ* hybridization light-micrographs, it is concluded that Npv30 transcripts accumulate in the infected cells of the central tissue (Fig. 8). These results and the finding that the expression of a chimeric *npv30-1* promoter-GUS gene was restricted to the infected cells in *Lotus corniculatus* transgenic nodules (Carsolio *et al.*, 1994) support the conclusion that Npv30 expression is restricted to the bean nodule infected cells.

Acknowledgments.

We gratefully acknowledge to Alejandra Covarrubias, Omar Pantoja and Marco Antonio Villanueva for critical reading of the manuscript. We also thank to Jose Luis Ortega for his help with dot blot analysis.

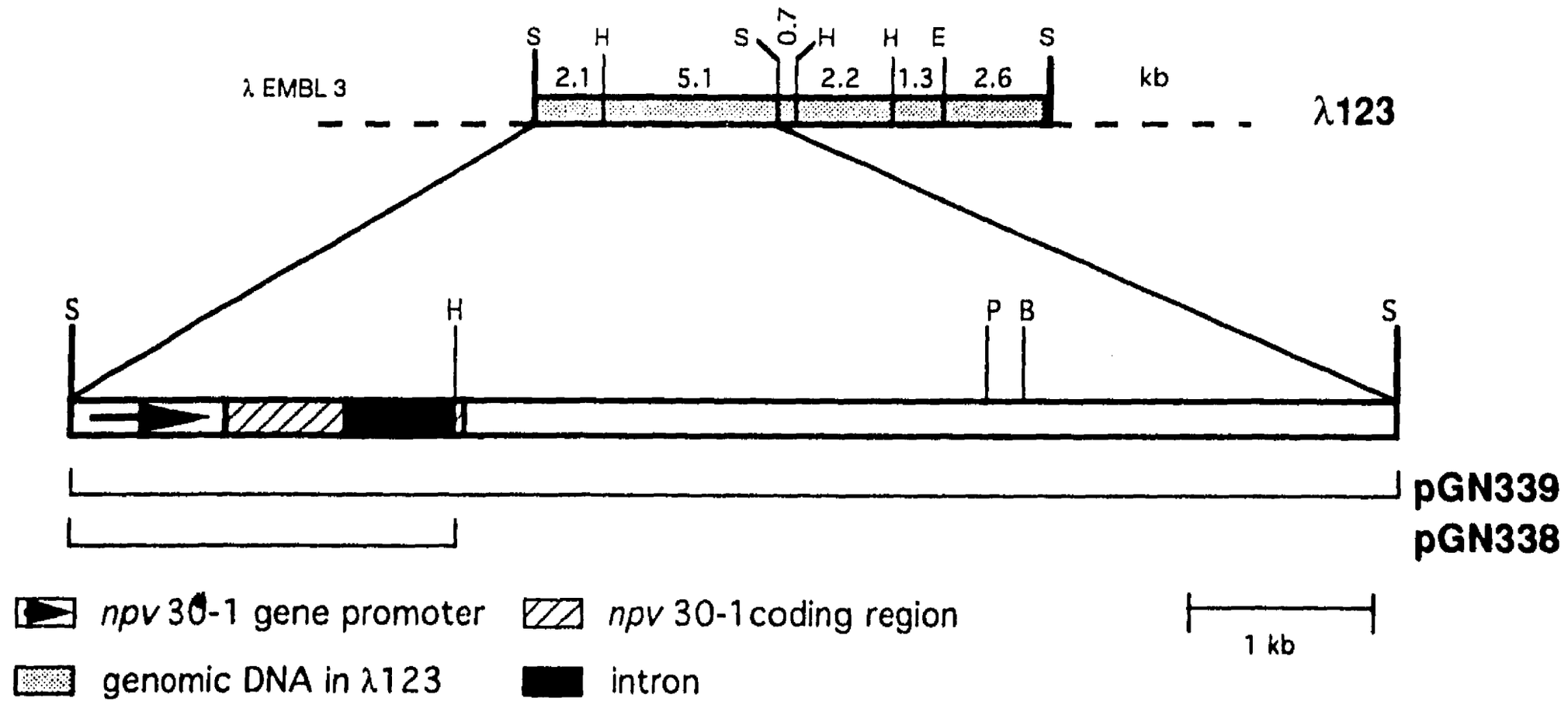

Figure 1. Restriction map of the genomic region of *P. vulgaris* present in the λ123 clone is displayed. The regions of *npv*-30 are diagrammed as it is shown in the figure. DNA fragments contained by the two derived plasmids, pGN338 and pGN339, used to determine the sequence of *npv30*-1, are represented with lines. The restriction sites are as follows: B= *Bam*HI, H= *Hind*III, E= *Eco*RI, P= *Pst*I, S= *Sal*I.

```
          Met Arg Ala Ile Leu Ile Thr Leu Phe Leu Ile Leu Ser Val Val Val Ala Glu Glu Ala
npv30-1   ATG AGA GCC ATA CTA ATT ACT CTG TTC TTG ATC CTA AGT GTG GTA GTT GCA GAA GAG GCA    60
N311                                                                      AG GCA     5

          Glu Asp Ala Ala Ile Val Glu Thr Ile Asp Pro Ala Lys Glu Ala Gly Ile Ser Val Ala
npv30-1   GAA GAT GCT GCA ATT GTT GAA ACC ATT GAT CCT GCA AAA GAA GCA GGA ATT TCT GTA GCA   120
N311      GAA GAT GCT GCA ATT GTT GAA ACC ATT GAT CCT GCA AAA GAA GCA GGA ATT TCT GTA ACA    65
                                                                                  Thr

          Thr Asn Pro Ala Lys Asp His Gly Ile Gly Gly Thr Gly Glu Ile Asn Asp Leu Ala Glu
npv30-1   ACT AAT CCT GCA AAA GAT CAT GGA ATT GGT GGA ACT GGT GAA ATC AAT GAT CTT GCT GAA   180
N311      ACT AAT CCT GCA AAA GAT CAT GGA ATT GGT GGA ACT GGT GAA ATC AAT GAT CTT GCT AAA   125
                                                                                  Lys

          Asp Ala Gly Val Gly Ile Ser Lys Ala Ile Tyr Gln Thr Leu Ser Gly Gln Pro Glu Ala
npv30-1   GAT GCT GGA GTT GGT ATT AGC AAA GCC ATT TAT CAA ACA CTT AGT GGG CAA CCT GAA GCG   240
N311      GAT GCT GGA GTT GGT ATT AGC AAA GCC ATT TAT CAA ACA CTT AGT GGG CAA CCT GAA GCG   185

          Tyr Glu Ser Pro Arg Phe Lys Arg Phe Val Thr His Cys Ser Ser His Val Ala Glu Thr
npv30-1   TAC GAA TCT CCA AGA TTC AAG AGG TTT GTG ACA CAT TGC AGC TCA CAT GTT GCT GAA ACA   300
N311      TAC GAA TCT CCA AGA TTC AAG AGG TTT GTG ACA CAT TGC AGC TCA CAT GTT GCT GAA ACA   245
          RsaI

          Cys Ser Asp Pro Met His Tyr Glu Gly Gly Ile Arg Asn Pro Thr Gly Leu Ser His Cys
npv30-1   TGC AGT GAT CCA ATG CAC TAT GAG GGT GGA ATC CGT AAC CCA ACT GGG TTG TCT CAC TGC   360
N311      TGC AGT GAT CCA ATG CAC TAT GAG GGT GGA ATC CGT AAC CCA ACT GGG TTG TCT CAC TGC   305

          Ile Phe Asp Ser Met Lys Ala Cys Leu Ala Asn His Lys Ala Ser Leu Tyr         Asp
npv30-1   ATT TTT GAT TCC ATG AAA GCA TGC TTG GCA AAT CAT AAA GCC TCG CTT TAT         GAC   414
N311      ATT TTT GAT TCC ATG GAA GCA TGC TTG GCA AAT CAT AAA GCC TCG TTT TCG TTT TAT GAC   365
                          Glu                                     Phe Ser Phe Tyr

          Ser Ala Arg Ser Lys Thr Leu Asn Leu Lys Pro Thr Lys Val Glu Tyr Leu Pro Val Ile
npv30-1   TCC GCT CGT TCC AAA ACC CTA AAT CTT AAA CCC ACA AAA GTC GAA TAT TTA CCG GTT ATC   474
N311      TCT GCT CGT TCC AAA ATC CTA AAC CTT AAA CCC ACA AAA GTC GAA TAT TTG CCG GTT ATC   425
                             Ile

          Ile Gln Thr Val Lys Phe Gln Thr Val Trp Lys Thr Cys Ser Gln Val Ser Ala Gln Ser
npv30-1   ATT CAG ACA GTA AAA TTT CAA ACT GTG TGG AAA ACC TGC TCT CAA GTC AGT GCA CAA AGT   534
N311      ATT CAG ACA GTA AAA TTT CAA ACT GTG TTG AAA ACC TGC TCT CAA GTC AGT GCA CAA AGT   485
                                            Leu

          Cys Leu Ser Asp Ser Asp Val Asp Ala Ser Thr Leu Gly Ala Cys Leu Leu Pro Ser Phe
npv30-1   TGT TTG AGT GAT TCT GAT GTT GAT GCA TCA ACT TTA GGA GCT TGT CTC TTA CCA TCT TTT   594
N311      TGT TTG AGT GAT TCT GAT GTT GAT GCA TCA ACT TTA GGA GCT TGT CTC TTA CCA TCT TTG   545
                                                                                  Leu

          Asn Gln Cys Val Tyr Pro Thr         Pro Pro         Pro Pro Pro Pro Pro Pro
npv30-1   AAC CAG TGT GTG TAT CCT ACT         CCA CCT         CCA CCA CCT CCA CCA CCT       639
N311      AAC CAG TGT GTG TAT CAT ACT CAA ATG CCA CCT ATA CCT CCA CCA CCT CCA CCA CCT CTA   605
                          His     Gln Met         Ile Pro                           Leu

                                                  Pro Pro     A
npv30-1                                           CCT CCT     G gtaagacttctttaataatata   668
N311      CGT CCA CCA CCT CAA CGT CCA CCA CCT CCT ACT CCA CCT GCT G                       648
          Arg Pro Pro Pro Gln Arg Pro Pro Pro Pro Thr         Ala

npv30-1   tatgatttaagaaatataattaacacactttttatttctgtattttgttaacaaaactcattttagattaacctctacaaa  748
npv30-1   attctacctttcaccactacttttttcattaaaacaaaaaaaaatataatattaaaaaatataatatttagtttagagtttat  828
npv30-1   ggtttagaattaaagattaaggtttacacatacaaactctaaattctagatcttatatcgatttaaaaataaaaatatttt  908
npv30-1   cataatatataaattttatgttttatttttattataaaattacataaatgcatcatttcatgtgaaatttttgaaaaacat  988
npv30-1   gtaaataatttcgaaaaccattacatgagcagatatattcataatggaaatagaagtagtttcatcaaacaaattaatat 1068
npv30-1   tataggattttcaaaaacataacactaatgtggcttcttatagtttttcttacattatgtgttcttattatatctgatg 1148
npv30-1   aatcacatgttttccattgcttcttctttttttctggaattaatgtttatattttagaagcttaaaaggttgtggtttgtg 1228

             sp Glu Thr Arg Arg Ter
npv30-1   cag AT GAG ACA CGA AGA TAA atcatagtggggaagtactt                                  1268
N311         AT AAG ACA AGA AGA TAA atcatactggg aagtacttatagttagcattgccatattatgctaatacaac  718
               Lys                                 PsaI

N311      aataatcgaatgctttctccatcatcttatatatgtaataaataatctctttttttctagctaagattaaactcttaatct  798

N311      cgttcttaatttaatttatggtgaaagactgttataccttc                                        838
```

Figure 2. Nucleotide and deduced amino acid sequences from the cDNA (pn311) and genomic (*npv30*-1) clones. Two potential polyadenylation signals (=) located in the cDNA and internal *Rsa*I restriction sites are also indicated. The arrow indicates the potential cleavage site of the putative signal peptide.

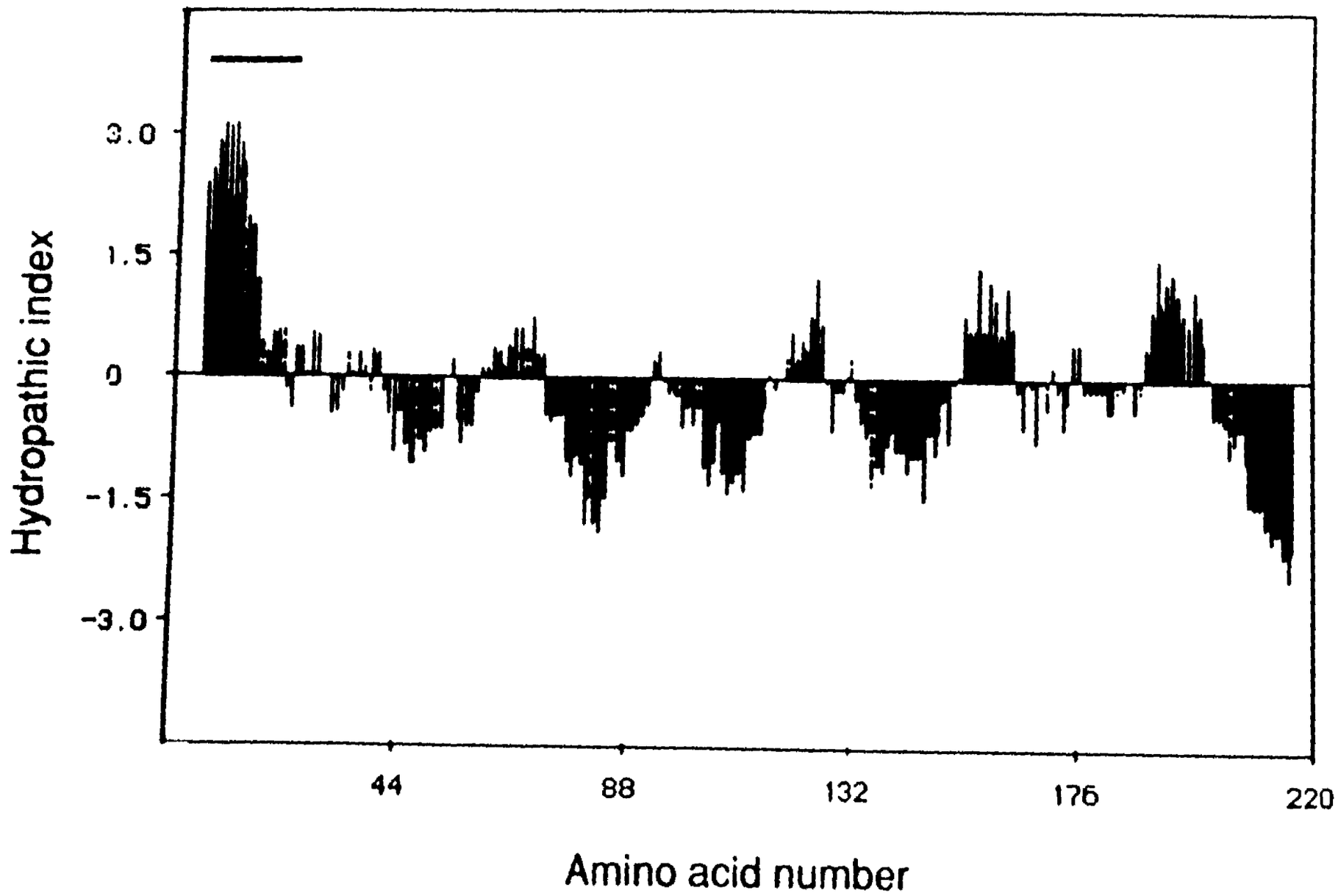

Figure 3. Hydropathic profile of the deduced Npv30-1 amino acid sequence determined by the method of Kyte and Doolittle (1982) with a window of 11 residues. The horizontal bar above indicates the region for the putative signal peptide.

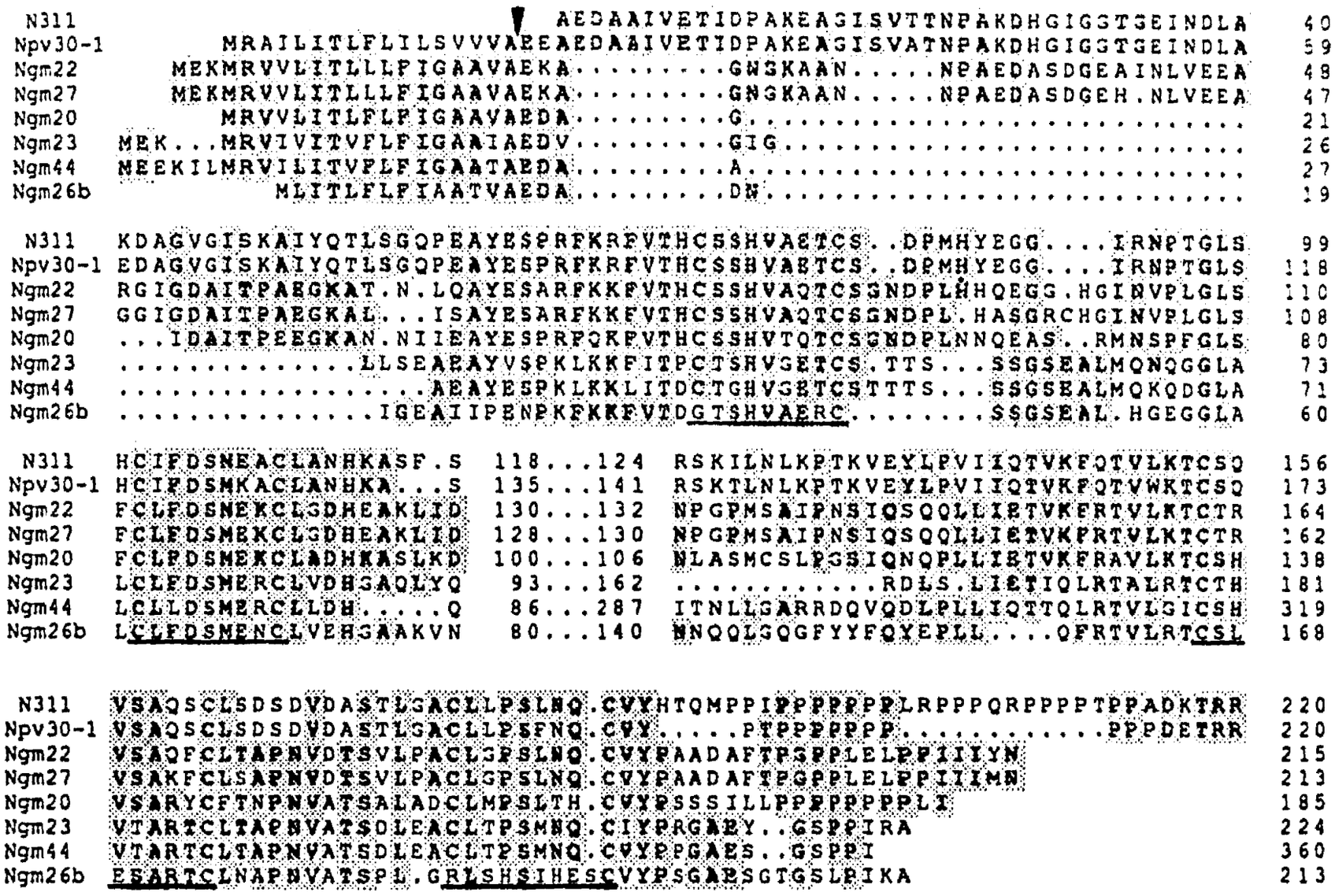

Figure 4. Comparison of the deduced amino acid sequence from the cDNA (pN311) and genomic (*npv30-1)* clones of Npv30 bean family, with members of soybean nodulin-A family; Ngm-44, Ngm-27, Ngm-26b, Ngm-23, Ngm-22 and Ngm-20. The shadowed areas show the conserved regions. The arrow indicates the potential cleavage site of the putative signal peptide.

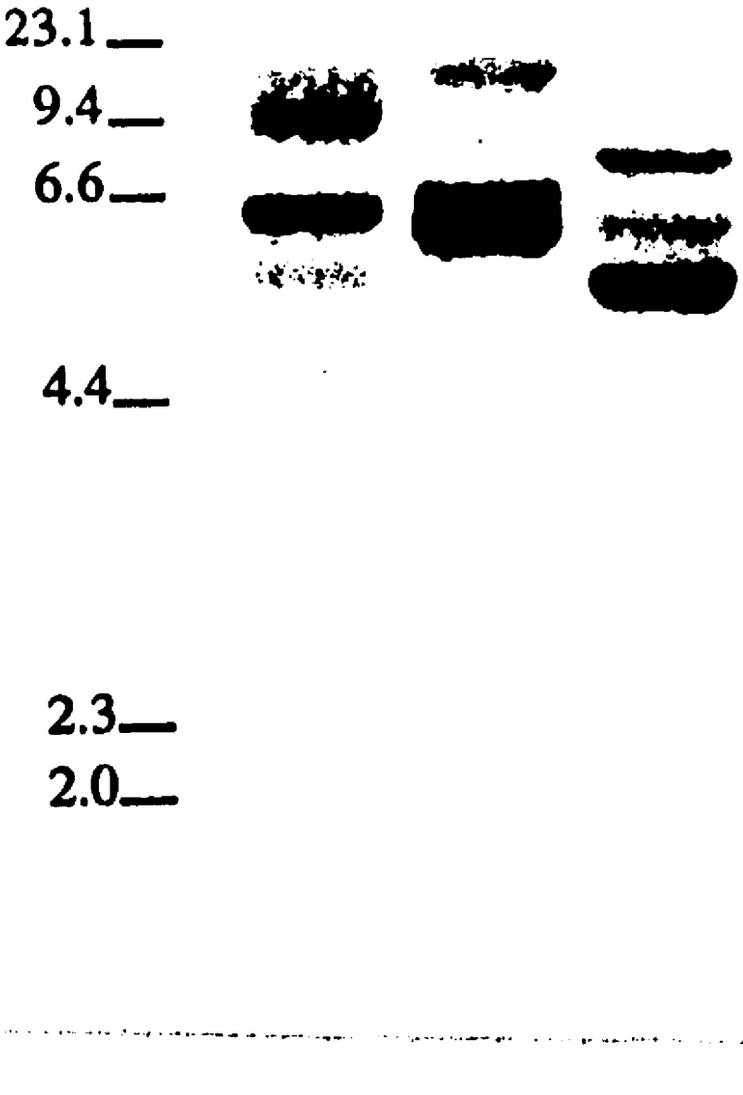

Figure 5. Genomic DNA Southern blot analysis of *Phaseolus vulgaris* L.Negro Jamapa. The DNA was digested with *Bgl*II (1), *Eco*RI (2) and *Hind*III (3). The blot was hybridized to the *Pst*I insert of pN311 cDNA clone. DNA markers are indicated in kb.

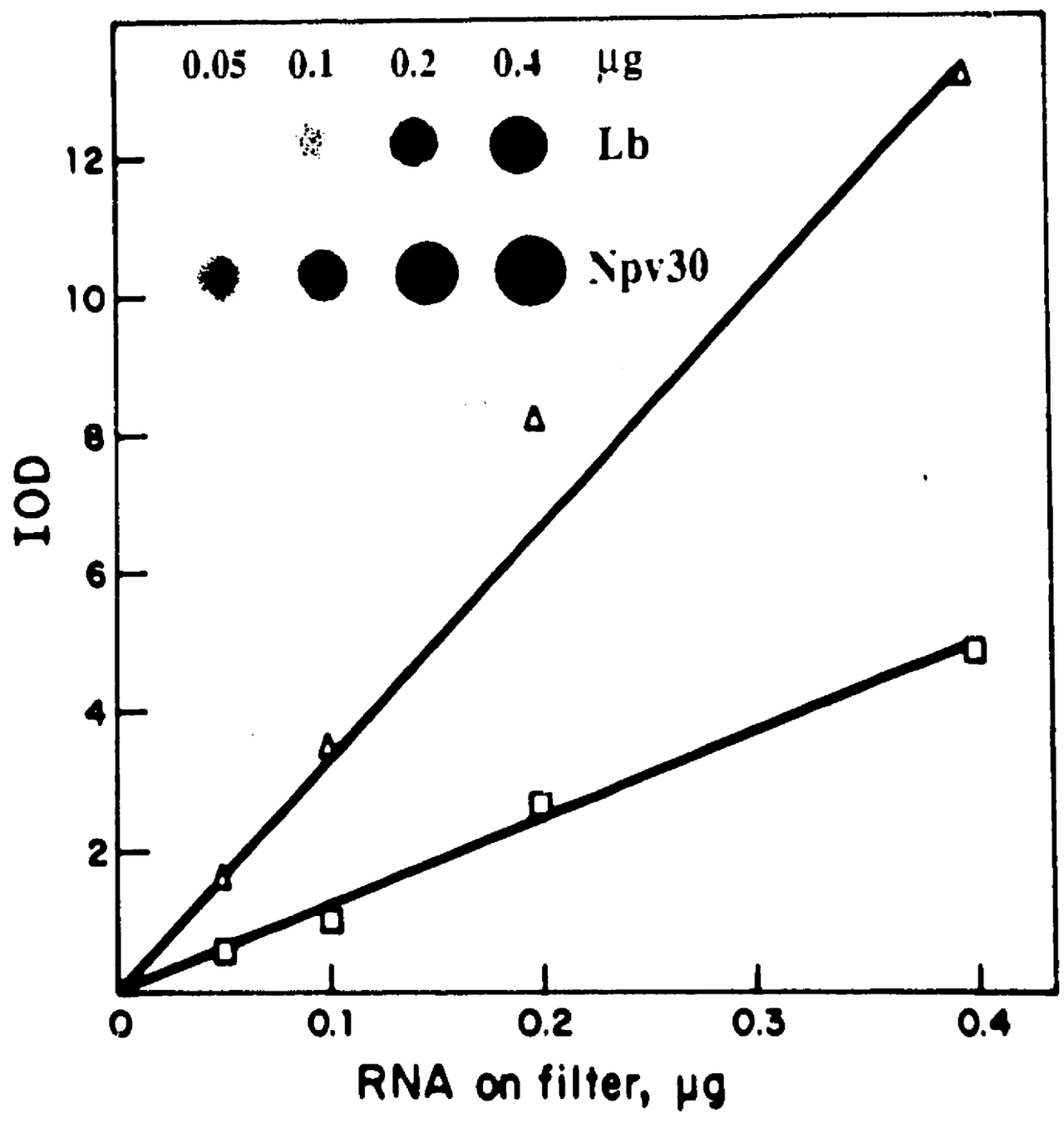

Figure 6. Determination of relative concentration of Npv30 transcripts compared to leghemoglobin mRNAs. Serial dilutions of total RNA from 21-d-old nodules were fixed to filters and hybridized to *Pst*I pN311 insert (Δ) and *Pst*I pNF-Lb01 insert (p). Autoradiography of dot blot hybridization is showed in the inset. The mRNA levels were quantified using an analytical imaging instrument and plotted versus µg of RNA.

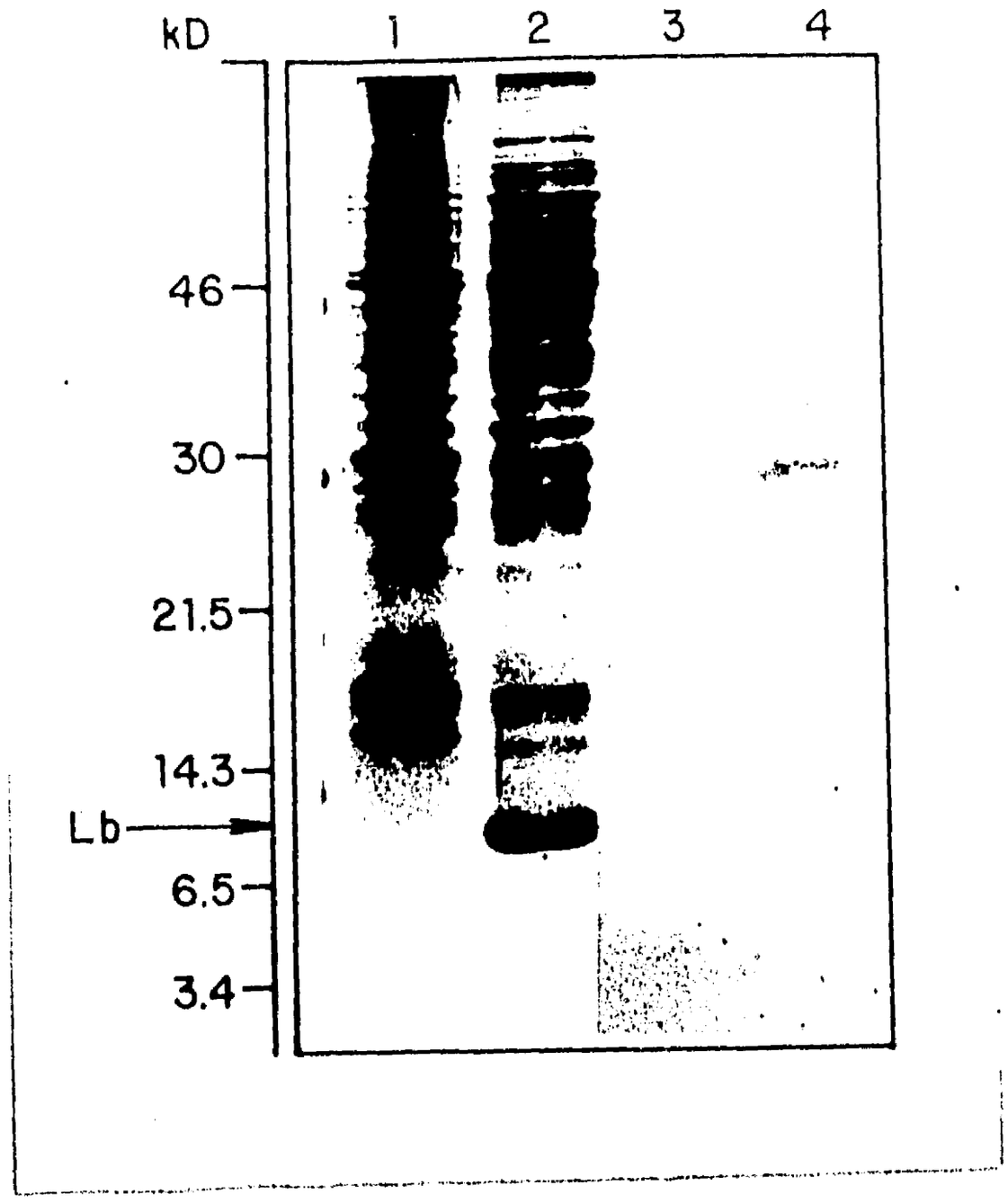

Figure 7. Western immunoblotting analysis of total 8-d-old uninfected root fraction (lanes 1 and 3), and total 21-d-old nodule fraction (lanes 2 and 4). Lanes 1 and 2 were stained with Coomassie blue, and lanes 3 and 4 were immunoblotted with antibodies to β-galactosidase-Npv30 fusion protein. Molecular weight markers are shown in kDa on the left margin. The arrow indicates the position where leghemoglobin migrated.

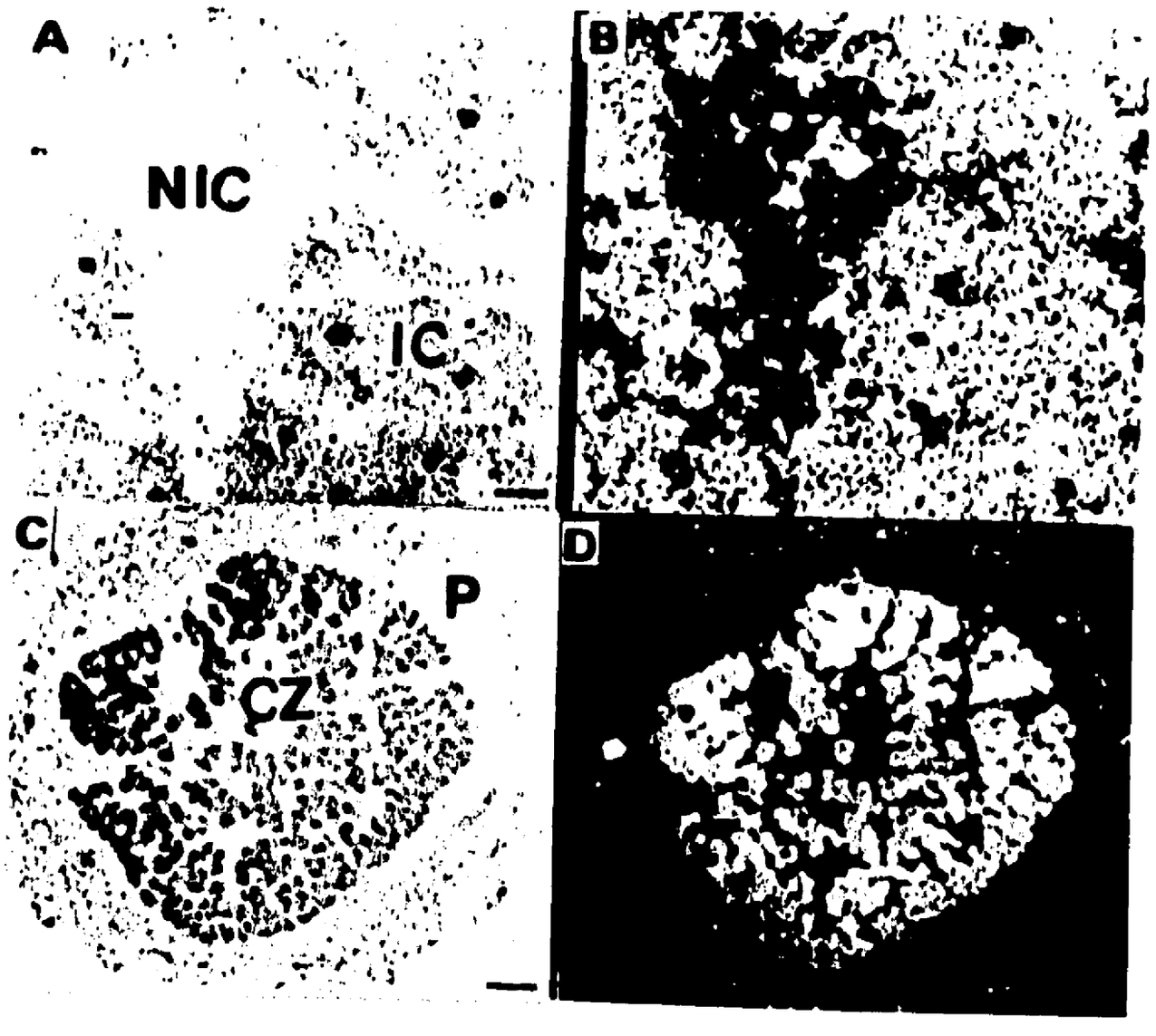

Figure 8. *In situ* localization of Npv30 transcripts in cross sections of 21-d-old nodules of *Phaseolus vulgaris* L. A, C are bright field micrographs; B, D are dark field micrographs in which silver grains are visible as white dots. The Npv30 cDNA here described was used to make a ^{35}S-UTP labeled antisense probe. A, B central tissue, bar equals 25 µm; C, D bar equals 500 µm. Abbreviations: (NIC) non-infected cell, (IC) infected cell, (CZ) central zone, (P) nodule parenchyma.

Literature Cited

Appleby CA (1984) Leghemoglobin and *Rhizobium* respiration. Annu Rev Plant Physiol 35: 443-478.

Berg JM (1990) Zinc finger domains: Hypothesis and current knowledge. Annu Rev Biophys Biophys Chem 19: 405-421

Bergmann H, Preddie E, Verma DPS (1983) Nodulin-35: A subunit of specific uricase (uricase II) localized in uninfected cells of nodules. EMBO J 2: 2333-2339.

Blake MS, Johnston KH, Russell-Jones J, Gotschlich EC (1984) A rapid sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on Western blots. Anal Biochem 136: 175-179

Campos F, Padilla JE, Vázquez M, Ortega JL, Enriquez C, Sánchez F (1987) Expression of a nodule-specific gene in *Phaseolus vulgaris* L. Plant Mol Biol 9: 521-532.

Carsolio C, Campos F, Sánchez F, Rocha-Sosa M (1994) The expression of a chimeric *Phaseolus vulgaris* nodulin 30-GUS gene is restricted to the rhizobially infected cells in transgenic *Lotus corniculatus* nodules. Plant Mol Biol 26: 1995-2001

Cooper JB, Long SR (1994) Morphogenetic rescue of *Rhizobium meliloti* nodulation mutants by trans-zeatin secretion. Plant Cell 6: 215-225.

Das A (1990) Overproduction of proteins in *Escherichia coli*. Methods Enzymol 182: 93-112.

de Vries S, Hoge H, Bisseling T (1991) Isolation of total and polysomal RNA from plant tissues. *In* SB Gelvin, RA

Schilperoot, DPS Verma, ed. Plant Molecular Biology Manual. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Dehio C, de Bruijn FJ(1993)The early nodulin gene *SrEnod2* from *Sesbania rostrata* is inducible by cytokinin. Plant J 2: 117-128.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12: 13-15.

Fisher RF, Long SR (1992) *Rhizobium*-plant signal exchange. Nature 357: 655-659.

Franssen HJ, Vijn I, Yang WC, Bisseling T (1992) Development aspects of the *Rhizobium*-legume symbiosis. Plant Mol Biol 19: 89-107.

Fuller F, Kunstner P, Nguyen T, Verma DPS (1983) Soybean nodulin genes: Analysis of cDNA clones reveals several major tissue-specific sequences in nitrogen-fixing root nodules. Proc Natl Acad Sci USA 80: 2594-2598.

Govers F, Nap JP, van Kammen A, Bisseling T (1987) Nodulins in the developing root nodule. Plant Physiol Biochem 25: 309-322.

Gottlob-McHugh SG, Johnson DA (1991) Detection of a subfamily of genes within the soybean nodulin-A multigene family. Can J Bot 69: 2663-2669.

Hirsch AM (1992) Developmental biology of legume nodulation. New Phytologist 122: 211-237.

Jacobs FA, Zhang M, Fortin MG, Verma DPS (1987) Several nodulins of soybean share structural domains but differ in their subcellular locations. Nucleic Acid Res 15: 1271-1280.

Journet EP, Pichon M, Dedieu A, de Billy F, Truchet G, Barker DG (1994) *Rhizobium meliloti* Nod factors elicit cell-specific transcription of the ENOD12 gene in transgenic alfalfa. Plant J 6: 241-249.

Kardailsky I, Yang W-C, Zalensky A, van Kammen A, Bisseling T (1993) The pea late nodulin gene PsNOD6 is homologous to the early nodulin genes PsENOD3/14 and is expressed after the leghaemoglobin genes. Plant Mol Biol 23: 1029-1037.

Katinakis P, Verma DPS (1985) Nodulin-24 gene of soybean codes for a peptide of the peribacteroid membrane and was generated by tandem duplication of a sequence resembling an insertion element. Proc Natl Acad Sci 82: 4157- 4161.

Kouchi H, Hata S (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. Mol Gen Genet 238: 106-119.

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157, 105-132.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Lara M, Cullimore JV, Lea PJ, Johnston AWB, Lamb JW (1983) Appearance of a novel form of plant glutamine synthetase during nodule development in *Phaseolus vulgaris* L. Planta 157: 254-258.

Lara M, Porta H, Padilla JE, Folch J, Sánchez F (1984) Heterogeneity of glutamine synthetase polypeptides in *Phaseolus vulgaris* L. Plant Physiol 76: 1019-1023.

Lee JS, Verma DPS (1984) Structure and chromosomal arrangement of leghemoglobin genes in kidney bean suggest divergence in soybean leghemoglobin gene loci following tetraploidization. EMBO J 3: 2745-2752.

Legocki RP, Verma DPS (1980) Identification of "nodule-specific" host proteins (nodulins) in soybean involved in the development of *Rhizobium*- legume symbiosis. Cell 153-163.

Manen J-F, Simon P, van Slooten J-C, Osteras M, Frutiger S, Hughes GJ. (1991) A nodulin specifically expressed in senescent nodule of winged bean is a protease inhibitor. Plant Cell 3: 259-270.

Martínez-Romero E, Segovia L, Martins-Mercante F, Franco AA, Graham P, Pardo MA (1991) *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena sp*. Int J Syst Bacteriol 41: 131-135.

Mauro VP, Nguyen T, Katinakis P, Verma DPS (1985) Primary structure of the soybean nodulin-23 gene and potential regulatory elements in the 5'-flanking regions of nodulin and leghemoglobin genes. Nucleic Acid Res 13: 239-249.

Nap JP, Bisseling T (1990) Developmental biology of plant-prokaryote symbiosis: the legume root nodule. Science 250: 948-954.

Padilla JE, Campos F, Conde V, Lara M, Sánchez F (1987)

Nodule-specific glutamine synthetase is expressed before the onset of nitrogen fixation in *Phaseolus vulgaris* L. Plant Mol Biol 9: 65-74.

Pichon M, Journet E-P, Dedieu A, de Billy F, Truchet B, Barker DG (1992) *Rhizobium meliloti* elicits transient expression of the early nodulin gene ENOD12 in the differentiating root epidermis of transgenic alfalfa. Plant Cell 4: 1199-1211.

Richter HE, Sandal NN, Marcker KA, Sengupta-Gopalan C (1991) Characterization and genomic organization of a highly expressed late nodulin gene subfamily in soybeans. Mol Gen Genet 229: 445-452.

Rhüter U, Müller-Hill B (1983) Easy identification of cDNA clones. EMBOJ 2: 1791-1794.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning. A laboratory manual. Second edition. Series CSH. Laboratory Press, Cold Spring Harbor, NY, USA.

Sandal NN, Bojsen K, Marcker KA (1987) A small family of nodule specific genes from soybean. Nucleic Acid Res 15: 1507-1519.

Sánchez F, Campos F, Padilla JE, Bonneville JM, Enriquez C, Caput D (1987) Purification, cDNA cloning and developmental expression of the nodule-specific uricase from *Phaseolus vulgaris* L. Plant Physiol 84: 1143-1147.

Sánchez F, Padilla JE, Perez H, Lara M (1991) Control of nodulin genes in root-nodule development and metabolism. Annu Rev Plant Physiol Plant Mol Biol 42: 507-528.

Sengupta-Gopalan C, Pitas JW, Thompson DV (1986) Expression of host genes during root nodule development in soybeans. Mol Gen Genet 203: 410-420.

Scheres B, van de Wiel C, Zalenski A, Horvath B, Spaink H, van Eck H, Zwartkruis F, Wolters A, Gloudeman T, van Kammen A, Bisseling T (1990) The ENOD12 gene product is involved in the infection process during the pea-*Rhizobium* interaction. Cell 60: 281-294.

Showalter AM (1993) Structure and function of plant cell wall proteins. Plant Cell 5: 9-23.

Towbin H, Staehelin T, Gordon T (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. Proc Natl Acad Sci USA 76: 4350-4354.

Thummler F, Verma DPS (1987) Nodulin-100 of soybean is the subunit of sucrose synthase regulated by the availability of free heme in nodules. J Biol Chem 362: 14730-14736.

van Kammen A (1984) Suggested nomenclature for plant genes involved in nodulation and symbiosis. Plant Mol Biol Rep 2: 43-45.

von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucl Acid Res 14: 4683-4690.

Yu H, Chen JK, Feng S, Dalgarno DC, Brauer AW, Schreiber SL (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 76: 933-945.

III. Discusión y Conclusiones

Hemos caracterizado una familia de nodulinas de 30 kDa cuyo transcrito se acumula abundantemente en frijol (Campos *et al.*, 1987; Campos *et al.*, 1995; Carsolio *et al.*, 1994) y que es similar a la familia-A de nodulinas de soya (Gottlob-Mchugh y Johnson, 1991; Jacobs *et al.*, 1987; Mauro *et al.*, 1985; Sandal *et al.*, 1987; Sengupta-Gopalan *et al.*, 1986). La secuencia de amninoácidos deducida de la secuencia nucleotídica de dos miembros de la familia (clonas pN311 y *npv30-1*), muestra que la Npv30 posee tres dominios: 1) un péptido señal en el extremo amino-terminal, 2) una sección rica en prolinas en el extremo carboxilo-terminal, y 3) dos regiones con el motivo Cys-X_{7AAS}-Cys. En esta sección se discutirán los posibles papeles de estas regiones para la función de la Npv30.

La presencia de un péptido señal en el extremo amino-terminal de la Npv30 sugiere que estas proteínas pueden estar asociadas a membranas y/o ser compartimentalizadas. A pesar de que en este trabajo no analizamos la fracción membranal del nódulo, hemos detectado cantidades similares de Npv30 en la fracción "total" y fracción "soluble" cuando usamos el anticuerpo dirigido contra la proteína de fusión β-galactosidasa-Npv30, este dato, junto con el de la localización de los transcritos de Npv30 en células infectadas, nos hace suponer que la Npv30 puede encontrarse en el espacio peribacteroidal y que durante el proceso de extracción la membrana peribacteroidal se rompe, liberando el contenido a la fracción soluble; algo similar ocurre con

la uricasa del nódulo, la cual ha sido localizada mediante técnicas histológicas dentro de los peroxisomas. Sin embargo, al fraccionar bioquímicamente el nódulo, la uricasa es detectada en la fracción soluble (Bergmann *et al.*, 1983; Sánchez *et al.*, 1987). Aunque hemos tratado de inmunolocalizar la Npv30 en secciones del nódulo, no se ha obtenido hasta ahora ningún resultado. Esto puede ser debido a: 1) un bajo título del anticuerpo dirigido contra la proteína de fusión; 2) la pérdida y/o enmascaramiento de los epítopes que el anticuerpo reconoce; 3) los bajos niveles del antígeno correspondiente.

Una característica que posee el grupo de nodulinas Npv30 de frijol es la presencia de dos dominios que contienen cuatro residuos de cisteínas arregladas en pares con el motivo Cys-X7-Cys. Ambos dominios en un mismo gen tienen una gran homología, lo cual sugiere que este par de regiones ha sido el resultado de una duplicación "ancestral" (ver Fig.4). Esto también sugiere que ambos dominios deben llevar a cabo funciones homólogas. Se ha visto que las proteínas con pares de cisteínas o histidinas pueden formar estructuras llamadas "dedos de zinc" (ver Fig.5). Estos dedos de zinc forman estructuras que unen metales, principalmente zinc. Algunas de las proteínas que poseen estas estructuras son factores de transcripción que se unen a DNA o RNA (p.ej. GAL4 en *Saccharomyces cerevisiae*, TFIIIA en ovocitos de rana). También se han encontrado enzimas (p.ej. proteinasas) con estas estructuras. Se han analizado las secuencias de varios dedos de zinc y se ha obtenido una secuencia consenso para estos dominios (Berg, 1990), sin

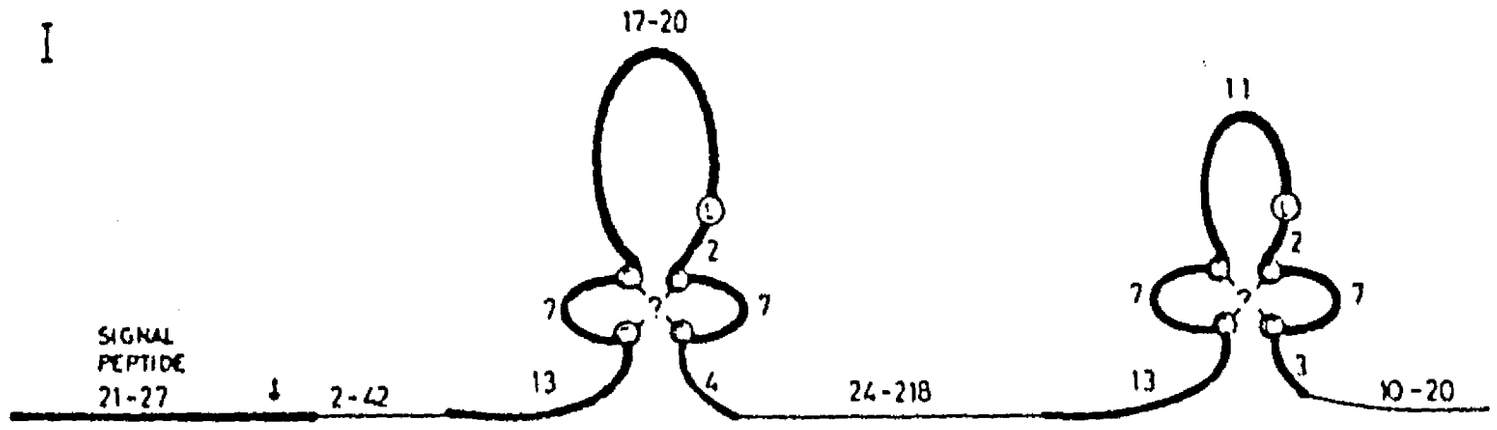

Fig. 5. **Estructura propuesta para los dos pares de cisteina de la familia Npv30 y la familia-A. I.** La línea gruesa muestra regiones conservadas, la línea delgada muestra regiones divergentes. **II.** Esquemas de la estructura de un dominio de unión a zinc. **A.** Estructura de un "dedo de zinc" de la enzima aspartato carbamoil transferasa de *E. coli*. **B.**Estructura del dominio de unión a zinc del factor IIIA de *Xenopus*. Basado de Sandal y cols., (1987).

embargo los pares de cisteínas que se encuentran en la familia Npv30 no tienen similitud con alguna secuencia tipo dedo de zinc reportada. Existe la posibilidad de que los pares de cisteínas que hay en estas nodulinas formen un nuevo grupo de dedos de zinc (Sandal *et al.*, 1987), aunque hasta ahora no se han realizado los experimentos para determinar el acoplamiento de metales por la Npv30. Otra posible función que pueden tener los pares de cisteínas de la Npv30, es la de establecer puentes disulfuro con otras proteínas o entre ellas mismas; resultaría interesante determinar el comportamiento de la proteína en geles nativos sin agentes reductores para analizar esta posibilidad.

Hemos encontrado que en las clonas analizadas de Npv30 hay una región rica en prolinas en el extremo carboxilo-terminal. Algunas de las funciones en las que podria estar involucrada esta región son: 1) causar de una vida media corta para la Npv30 (Rogers *et al.*, 1984); 2) formar dominios de interacción proteína-proteína (Yu *et al.*, 1994); 3) servir de punto de interacción con la pared célular vegetal, en caso de que los residuos de prolina se hidroxilen (Showalter, 1993). Es necesario llevar a cabo una serie de experimentos para elucidar el papel que tiene esta región en la función de la proteína. Sin embargo, aunque los dos miembros de la Npv30 hasta ahora analizados poseen esta región rica en prolina, sólo dos miembros de la familia-A de soya poseen una región similar (ver Fig. 4). De estos datos es posible pensar que la región rica en prolinas no es indispensable para la función de la proteína, aunque podría ser que esta región solo sea importante en frijol. Para

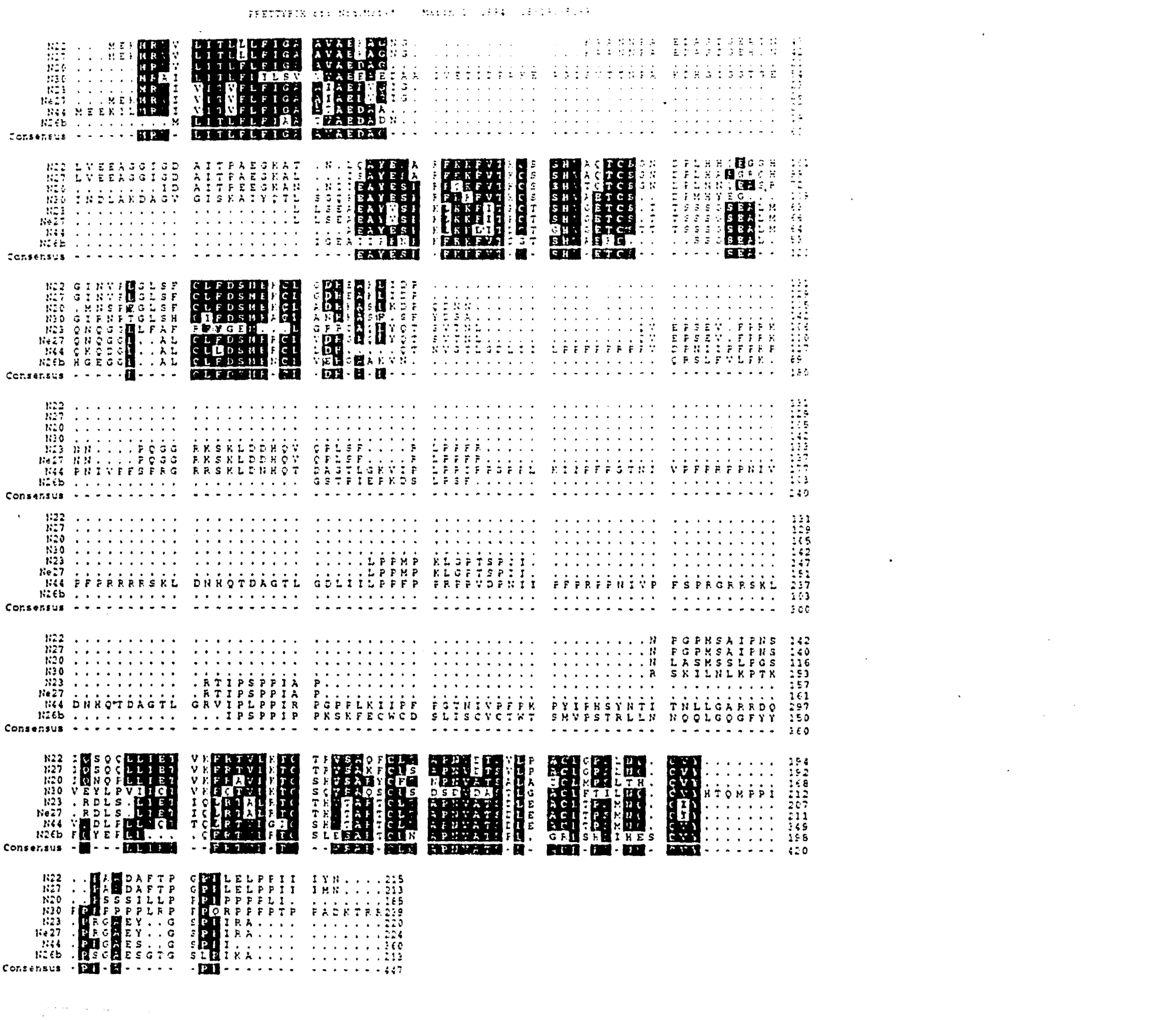

Fig. 4. **Alineamiento múltiple entre los miembros de la familia-A de nodulinas de soya y la Npv30 de frijol.** Las cajas negras indican aminoácidos identicos, las cajas grises indican cambios conservativos.

establecer si este dominio es una característica común de la familia de Npv30 en frijol es necesaria la secuenciación de mas miembros de la familia de Npv30.

Aunque los RNAs mensajeros que codifican para este grupo de nodulinas son los más abundantes del nódulo de frijol, los niveles de proteína que hay en el nódulo no parecen reflejar esta situación; la hipótesis que tenemos para explicar este fenómeno es que las Npv30s son proteínas de una vida media corta y que se degradan rápidamente en el nódulo; la sección rica en prolinas en el extremo carboxilo de esta familia podría estar involucrada en esto. Rogers *et al.*, (1984) han propuesto que la abundancia de prolinas en una proteína puede causar que ésta tenga una vida media corta.

Un hecho que nos ha intrigado mucho es el comportamiento *sui generis* que muestra este grupo de nodulinas cuando se analizan en geles bidimensionales (Campos *et al.*, 1987), ya que las Npv30s se visualizan como un barrido en la primera dimensión, atravesando el gradiente de pH (4.0-7.0). Suponemos que este comportamiento es debido a que la familia de Npv30 codifica para polipéptidos con un peso molecular semejante (30 kDa) pero con una heterogeneidad en la cola de prolinas en el extremo carboxilo-terminal, las cuales causarían un punto isoeléctrico aberrante; en los miembros analizados de esta familia hemos encontrado que uno de ellos tiene 10 prolinas sucesivas, mientras que el otro codifica para 18 prolinas en sus últimos 30 residuos del extremo carboxilo.

Al inicio de este Proyecto, suponíamos que al analizar la secuencia nucleotídica de estas nodulinas podríamos obtener indicios de su función. Sin embargo, al llevar a cabo el análisis, encontramos que esta familia de genes tiene similitud con una familia de nodulinas de soya con función desconocida (Gottlob-Mchugh y Johnson, 1991). Esta familia de genes ha sido designada como la familia-A (Jacobs *et al.*, 1987).

En las clonas genómicas analizadas tanto en frijol como en soya, se ha localizado un intrón cerca del extremo 3' no traducido, esto implica que tanto la familia Npv30 como la familia-A de soya derivan de un gen ancestral común.

Hasta ahora no se ha reportado la existencia de nodulinas similares en otras leguminosas además de frijol (este trabajo), soya (Gottlob-Mchugh y Johnson, 1991) y *Vigna* (Sengupta-Gopalan, comun. personal). Hemos tratado de detectar genes similares a Npv30 por medio de hibridización cruzada, en otros organismos, usando como sonda el cDNA de la Npv30. Aunque hemos detectado señal en RNA de nódulos de chícharo y en DNA de la levadura *S. cerevisiae*, estos resultados son muy preliminares y cabe la posibilidad de que se trate de hibridación inespecífica debido al alto contenido de CC de nuestra sonda.

Mediante hibridización *in situ* se han detectado los transcritos de Npv30 en las células infectadas del nódulo. Suponiendo que el producto de estos transcritos también se localice en estas células, podríamos suponer que la Npv30 apoye la función de las células infectadas (p.ej, asimilación de amonio, regulación de O_2), aunque habrá que

determinar la localización exacta de la Npv30 en el nódulo. Incluso la presencia del péptido señal en la Npv30 nos da indicios de que esta proteína pudiera estar asociada a la membrana peribacteroidal.

En conclusión, en este trabajo se ha demostrado que el grupo de nodulinas Npv30 son codificadas por una familia de genes similar a una familia de nodulinas de soya, cuyos transcritos se acumulan abundantemente en células infectadas del nódulo de frijol. El producto de estos genes es poco abundante.

A. Perspectivas

A pesar de haber realizado una caracterización mas exhaustiva de esta familia de genes, aún somos incapaces de postular alguna posible función para esta familia de nodulinas, suponemos que la estrategia usada hasta ahora no ha sido la más adecuada.

La obtención de mutantes en estos genes y el análisis de las mutantes seria la forma idonea para determinar la función de la Npv30, pero hasta ahora, esto es más que imposible de hacer en frijol. Una opción más plausible es la expresión o sobreexpresión de estos genes en otra planta (p.ej. tabaco o *Arabidopsis*) y la búsqueda de algun fenotipo (letalidad, efectos en el crecimiento, etc.).

Un enfoque que estamos aplicando es la expresión de la Npv30 en levadura; datos preliminares muestran que *S. cerevisiae* al expresar la Npv30 es más resistente a choque osmótico (Olivares JE, Tesis de Licenciatura).

IV. REFERENCIAS.

Appleby CA. (1984). Annu Rev Plant Physiol 35, 443.

Atkins CA, Pate JS, Sharkey PJ. (1975). Plant Physiol 56, 807.

Awonaike KO, Lea PJ, Miflin BJ. (1981). Plant Sci Lett 23, 189.

Bauer WD. (1981). Annu Rev Plant Physiol 32, 407.

Berg JM. (1990), Annu Rev Biophys Biophys Chem 19, 405.

Bergmann H, Preddie E, Verma DPS. (1983). EMBO J 2, 2333.

Bisseling T, Franssen H, Govers F, Horvath B, Moerman M, Scheres B, van de Wiel, Yang W-C. (1991). In: Molecular genetics of plant-microbe interactions, ed. Hennecke H, Verma DPS, Dordrecht, Boston, London, Kluwer Academic Publishers, pp.300-303.

Boland MJ, Benny G. (1977). Eur J Biochem 79, 355.

Brewin NJ. (1991). Annu Rev Cell Biol 7, 191.

Caetano-Anolles G, Crist-Estes DK, Bauer WD. (1988). J Bacteriol 170, 3164.

Campos F, Padilla JE, Vázquez M, Ortega JL, Enriquez C, Sánchez F. (1987), Plant Mol Biol 9, 521.

Cullimore JV, Lara M, Lea PJ, Miflin BJ. (1983). Planta 157, 245.

Currier WM, Sorbel GA. (1974). Science 146, 434.

de Bruijn FJ, Downie JA. (1991). Curr Opinion Biotech 2, 184.

Denarie J, Truchet G, Prome J-C. (1994). In: Advances in molecular genetics of plant-microbe interactions, eds. Daniels M, Downie JA, Osbourn AE, Dordrecht, Boston, London, Kluwer Academic Publishers, pp. 81-90.

Djordjevic MA, Redmond JW, Batley M, Rolfe BG. (1987). EMBO J 2, 1173.

Fahraeus G, Ljunggren H. (1959). Physiol Plant 12, 145.

Franssen HJ, Vijn I, Yang WC, Bisseling. (1992). Plant Mol Biol 19, 89.

Fuller F, Kunstner P, Nguyen T, Verma DPS. (1983). Proc Natl Acad Sci USA 80, 2594.

Gottlob-McHugh SG, Johnson DA. (1991). Can J Bot 69, 2663.

Govers S, Gloudemans T, Moerman M, van Kammen A, Bisseling T, (1985). EMBO J 4, 861.

Graham PH. (1963). J Gen Microbiol, 30, 245.

Hanks JF, Schubert KR, Tolbert NE. (1983). Plant Physiol 71, 869.

Hanks JF, Tolbert NE, Schubert KR. (1981). Plant Physiol 68, 65.

Hirsch AM. (1992). New Phytol 122, 211.

Kondorosi A. (1991). In: Molecular genetics of plant-microbe interactions, ed. Hennecke H, Verma DPS. Dordrecht, Boston, London, Kluwer Academic Publishers, pp. 11-118.

Kouchi H, Hata S. (1993). Mol Gen Genet 238, 106.

Lang-Unnasch N, Ausubel FM. (1985). Plant Physiol 77, 833.

Lara M, Cullimore JV, Lea PJ, Miflin BJ, Johnston AWB, Lamb JW. (1983). Planta 157, 254.

Lara M, Porta H, Padilla J, Folch J, Sánchez F. (1984). Plant Physiol 76, 1019.

Lea PJ, Awonaike O, Cullimore JV, Miflin BJ. (1982). Israel J Botany 31, 140.

Legocki RP, Verma DPS. (1979). Science 205, 190.

Legocky RP, Verma DPS. (1980). Cell 20, 153.

Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Prome JC, Denarie J. (1990). Nature 344, 781.

Long SR, Cooper J. (1988). In: Molecular genetics of plant-microbe interactions, ed. Palacios R, Verma DPS, Minnesota APS Press, 1988, pp. 163-178.

Long SR. (1989). Cell 56, 203.

Mauro VP, Nguyen T, Katinakis P, Verma DPS. (1985). Nucleic Acid Res 13, 329.

Mcparland RH, Guevara JG, Becker RR, Evans HJ. (1976). Biochem J, 153, 597.

Mellor RB, Christensen TMIE, Werner. (1986). Proc Natl Acad Sci USA 83, 654.

Mergaet P, van Montagu M, Prome J-C, Holters M. (1993) 90, 1551.

Miflin BJ, Lea PJ. (1980). In: The biochemistry of plants, vol. 5, ed. BJ Miflin, pp. 65-113, London, Academic Press.

Nap J-P, Bisseling. (1990). Science 250, 948.

Napoli CA, Hubell DH. (1975). Appl Microbiol 30, 1003.

Newcomb EM, Tandom SR. (1981). Science 212, 1394.

O'Farrell PH. (1975) J Biol Chem 250, 4007.

Ohyma T, Kumazawa K. (1980). Soil Sci Plant Nutr 26, 109.

Ortega JL, Campos F, Sánchez F. Lara M. (1986). Plant Physiol 80, 1051.

Padilla JE, Campos F, Conde V, Lara M, Sánchez F. (1987). Plant Mol Biol 9, 65.

Pate JS, Gunning BES, Brianty LG. (1969). Planta 85, 11.

Price NPJ, Relic B, Talmont E. (1992), Mol Microbiol 6, 3575.

Reynolds PHS, Boland MJ, Belvins DG, Schubert KR, Randall DD. (1982). Plant Physiol 69, 1334.

Robertson JG, Farnden KJF. (1980). In: The biochemistry of plants, vol. 5, ed. Miflin BJ, pp. 65-113, London, Academic Press.

Robertson JG, Lyttleton P, Bullivant S, Grayston GF. (1978a). J Cell Sci 30, 129.

Robertson JG, Warburton MP, Lyttleton P, Foryce AM, Bullivant S. (1978b). J Cell Sci 30, 151.

Robertson JG, Warburton PW, Farnden KJF. (1975). FEBS Lett 55, 33.

Roche P, Lerouge P, Prome JC, Faucher C, Vasse J, Maillet F, Camut S, de Billy F, Denarie J, Truchet G. (1990). In: Molecular genetics of plat microbe interactions, ed. Hennecke H, Verma DPS. Dordrecht, Boston, London, Kluwer Academic Publishers, pp, 119-126.

Rogers S, Wells R, Rechsteiner M. (1986), Science 234, 364.

Rolfe BG. (1988). Biofactors 1, 3.

Sánchez F, Campos F, Padilla JE, Bonneville JM, Enriquez C, Caput D. (1987). Plant Physiol 84, 1143.

Sánchez F, Padilla JE, Perez H, Lara M. (1991). Annu Rev Plant Physiol Plant Mol Biol 42, 507.

Sandal NN, Bojsen K, Marcker KA. (1987). Nucleic Acid Res 15, 1507.

Sanjuan J, Carlson RW, Spaink HP. (1992). Proc Natl Acad Sci USA 89, 8789.

Scheres B, van de Wiel, Zalenski A, Horvath B, Spaink H, van Eck, Zwartkruis F, Walters A, Gloudeman T, van Kammen A, Bisseling T. (1990).Cell: 281.

Schubert KR. (1981). Plant Physiol 67, 79s.

Schultze M, Quillet-Sire B, Kondorosi E. (1992). Proc Natl Acad Sci USA 89, 192.

Sen D, Schulman HM. (1980) New Phytol 85, 243.

Sengupta-Gopaland C, Pitas JW, Thomson DV, Hoffman LM. (1986) Mol Gen Genet 203, 410.

Showalter AM. (1993). Plant Cell 5, 9.

Spaink HP. (1992). Plant Mol Biol 20, 977.

Sprent JI: (1989). New Phytol 111, 129.

Thumler F, Verma DPS. (1987). J Biol Chem 362, 14730.

van Egerat AWSM. (1972). Meded Land Wagenin 72.

van Kammen. (1984). Plant Mol Biol Rep 2, 43.

van Kammen. (1984). Plant Mol Biol Rep 2, 43.

Verma DPS, Kazazian V, Zogbi V, Bal AK. (1978b).

Verma DPS, Zogbi V, Bal AK. (1978a). Plant Sci Lett 13, 137.

Vincent JM.(1980). In: Nitrogen Fixation, ed. Newton WE, Orme-Johnson WH, vol.2, pp. 103-122. Baltimore, University Park Press.

Werner D. (1992). Symbiosis of plant and microbes. Published by Chapman & Hall, London.

Yu H, Chen JK, Feng S, Dalgarno DC, Brauer AW.(1994). Cell 76, 933.