Localization of functional regions of the Rhizobium nodD product using hybrid nodD genes

Spaink, H.P.; Wijffelman, C.A.; Okker, R.J.H.; Lugtenberg, E.J.J.

Plant Molecular Biology 12: 59–73 (1989)

Localization of functional regions of the *Rhizobium nodD* product using hybrid *nodD* genes

Herman P. Spaink, Carel A. Wijffelman, Robert J. H. Okker* and Ben E. J. Lugtenberg
*Department of Plant Molecular Biology, Leiden University, Nonnensteeg 3, 2311 VJ Leiden, Netherlands (*author for correspondence)*

Received 5 July 1988; accepted in revised form 12 October 1988

Key words: autoregulation, flavonoid independent transcription, flavonoid specificity, gene regulation, *nodD* gene, *Rhizobium*

Abstract

The flavonoid-inducible *nod* promoters of *Rhizobium* are positively regulated by the *nodD* gene which is highly conserved in various *Rhizobium* species. The *nodD* genes are functionally different in (i) their response to various exogenously added flavonoid inducers, (ii) the extent to which they mediate the activation of the flavonoid-inducible promoters, and (iii) the extent to which they repress their own constitutive transcription. In order to localize the regions of the *nodD* product which determine these differences, two series of *nodD* hybrid genes have been constructed. In one series the 5′ moiety is derived from the *R. meliloti nodD1* gene and the 3′ moiety from the *R. trifolii nodD* gene. In the other series, the origins of the *nodD* moieties are reversed. Two regions of the *nodD* product appeared to be involved in autoregulation and it was also shown that the *nodD* promoters differ in their susceptibility to autoregulation. Many regions, dispersed over the entire *nodD* product, are involved in the specificity of activation by flavonoids. Several hybrid *nodD* genes were characterized which activate transcription with novel inducers. Furthermore, two classes of hybrid *nodD* genes were found from which the activation characteristics differ completely from those of the parental *nodD* genes. The first class activates the *nodABCIJ* promoter to the maximum level in the absence of flavonoid inducer. This level can no longer be influenced, positively or negatively, by the presence of (iso-)flavonoids. With the second class of hybrids, activation of the *nodABCIJ* promoter, even in the presence of flavonoid inducers, is no longer possible.

Introduction

Bacteria of the genus *Rhizobium* are able to establish a symbiosis with leguminous plants resulting in the formation of root nodules in which atmospheric nitrogen is fixed. Each bacterial species is restricted to its specific group of host plants (cross-inoculation group). Several nodulation (*nod*) genes, designated *nodFE, nodG* and *nodH*, determining this host range, have recently been identified in at least one of the fast-growing *Rhizobium* species [4, 31, 16, 2, 35].

In contrast to these so-called host-specific *nod* genes, the five *nod* genes *ABCIJ,* which constitute one operon, are functionally interchangeable, i.e. common, between *Rhizobium* species (for a review see [7]). Both common and host-specific *nod* genes, which are present on large Sym(biosis) plasmids, are thought to be regulated at the transcriptional level as one regulon in which the constitutively expressed *nodD* product acts as a positive regulator [23, 18, 26, 31, 32]. This *nod* regulon is only transcribed in the presence of an inducer exuded by the roots of the

host plants. The inducing compounds found in these exudates have been identified as flavones, flavanones or closely related compounds [24, 25, 9, 37, 40]. A strongly conserved regulatory domain, designated the *nod* box, is present upstream for all inducible *nod* operons and appears to be essential for the inducible promoter activity [27, 30, 33].

The *nodD* products of the fast-growing *Rhizobium* species *R. leguminosarum, R. trifolii,* and *R. meliloti,* the last of which contains three copies of the *nodD* gene [11, 13], are highly homologous [8, 31, 30] and are common in their ability to activate heterologous inducible *nod* promoters [34]. However, they are strikingly different in a number of other properties.

(i) Their response to various flavonoids or exudates is different, a feature which has been shown to be the basis of the host-specific properties of the *nodD* gene [17, 34].

(ii) In otherwise isogenic bacteria the maximal level of *nod* gene induction is substantially lower in the presence of the *R. meliloti nodD1* gene then in the presence of the *R. leguminosarum* or *R. trifolii nodD* gene.

(iii) The *nodD* product represses its own transcription in the case of *R. leguminosarum* [26, 32] whereas with the *R. meliloti nodD1* gene this so-called autoregulation could not be demonstrated [23].

Little is known about the relationship between the predicted primary protein structures encoded by the *nodD* genes and the properties of these proteins. Using random mutagenesis, Burn *et al.* [1] have obtained four classes of mutations in the *nodD* gene of *R. leguminosarum* which affect autoregulation and/or transcriptional activation. One of these mutations has been localized by nucleotide sequencing. Horvath *et al.* [17] used a different approach by constructing a hybrid *nodD* gene consisting of the 5′ terminal part of the broad host range *Rhizobium* MPIK3030 *nodD1* gene and the *R. meliloti nodD1* 3′ terminal part using the conserved *Bam* HI restriction site common for both *nodD* genes. Since the latter approach depends on the very limited number of identical restriction sites in two different *nodD* genes, we have used a more general method to construct hybrid genes, i.e., by *in vivo* homologous recombination between *nodD* genes, in a manner analogous to the procedure used by Tommassen *et al.* [36]. Characterization of the hybrid proteins both with respect to fusion site and properties in which the parent *nodD* proteins differ will then allow the localization of these properties to particular regions of the respective genes.

We have made hybrids between the *nodD1* gene of *R. meliloti* and the *nodD* gene of *R. trifolii.* Through phenotypical and genotypical characterization of the constructed hybrid gene products, we have localized several regions of the *nodD* product which are involved in autoregulation and in the specificity of activation by flavonoids. Furthermore, several types of *nodD* hybrids were identified whose characteristics differ strongly from those of the parent *nodD* genes.

Materials and methods

Bacterial strains and plasmids

Rhizobium strain LPR5045 is a Sym plasmid-less, rifampicin-resistant derivative of *R. trifolii* strain RCR5 [15]. *Escherichia coli* strains were JM101 [39] for transformation of bacteriophage M13 derivatives and KMBL1164 (*E. coli* K12 *del(lac-proAB) thi* F^-) for all other purposes. Broad host range plasmids were mobilized from *E. coli* to *Rhizobium* using pRK2013 as a helper plasmid [3]. Selection of transconjugants was done on YMB medium [14] with the addition of 10 mg l^{-1} chloramphenicol and 1 g l^{-1} streptomycin (with IncQ plasmids) or 2 mg l^{-1} tetracycline (with IncP plasmids) for plasmid selection and 20 mg l^{-1} rifampicin to select against *E. coli.* To measure transcriptional activation of the flavonoid inducible *R. leguminosarum nodABCIJ* promoter, IncP plasmids containing *nodD* genes or hybrid *nodD* genes were mobilized to the strain LPR5045·pMP154 [34]. Plasmid pMP154 is an IncQ transcriptional fusion plasmid, containing the *nodA* promoter of *R. leguminosarum* Sym plasmid pRL1JI cloned as a 114 bp restriction fragment in front of the *E. coli lacZ* gene.

DNA manipulation

DNA manipulation and transformation techniques were essentially according to Maniatis *et al.* [21]. Lyophilized large fragment (Klenow) of DNA polymerase I was obtained from BRL (Gaithersburg, MD, USA) and other enzymes were obtained from Boehringer (Mannheim, FRG).

Construction of hybrid nodD *genes by* in vivo *recombination*

To isolate hybrids resulting from *in vivo* homologous recombination, 5 μg of plasmid pMP600 or pMP800 DNA was treated with 80 units of *Eco* RI for 3 hours and subsequently transformed to *E. coli* KMBL1164. The insertion of the CAT gene in the used plasmids is an improvement of the method described by Tommassen *et al.* [36], used to construct hybrid genes of *ompC* and *phoE* of *E. coli*, since the presence of this extra marker can be used to detect transformants which resulted of an (unavoidable) incomplete restriction endonuclease treatment. To minimize the occurrence of clones resulting from the same cross-over, several DNA preparations of pMP600 and pMP800 have been used, isolated from independent cultures grown in the absence of chloramphenicol. These cultures were taken from a bacterial culture which was grown continuously in the presence of chloramphenicol and in which therefore no hybrid *nodD*-containing plasmids were present. The frequency in which chloramphenicol-sensitive colonies were obtained with the used DNA preparations varied between 5% and 20%. All plasmids of chloramphenicol-sensitive colonies were analyzed using *Hind* III and *Bam* HI restriction sites, the positions of which are indicated in Fig. 1B. In this screening we found that 10% of plasmids apparently did not result from homologous recombination.

Nucleotide sequence determination

DNA sequencing was performed by the dideoxy chain termination method [28] using the M13 vectors tg130 or tg131 [19]. Labeled (α-^{35}S)dATP was purchased from Amersham International plc (Amersham, UK). Restriction sites used for the cloning in M13 vectors were *Bgl* II, *Bam* HI, *Pvu* II, *Ava* II and *Sma* I.

Analysis of nod *gene expression*

Expression of the *nodABCIJ* promoter of *R. leguminosarum,* indicated by the level of produced β-galactosidase, was measured according to methods described previously [40]. Expression of the *nodD* promoter was measured in an identical way. Djordjevic *et al.* [6] have shown that the expression of the *nodD* gene of *R. trifolii* is also dependent on the growth phase of the bacteria. Our experiments have therefore been performed under standardized growth conditions using cells in the logarithmic growth phase (OD 660 = 0.1 which corresponds to 2×10^8 bacteria per ml). Assays of β-galactosidase were carried out in triplicate and were reproducible within 20% from experiment to experiment. If applicable, flavonoid inducers were added to a final concentration of 400 nM. Luteolin was obtained from Carl Roth (Karlsruhe, FRG), 7-hydroxyflavone was from EGA (Steinheim, FRG), daidzein and genistein were from Sarsyntex (Merignac, France), hesperitin was from Fluka AG (Buchs, Switzerland) and umbelliferone was from Sigma (St Louis, MO, USA).

Results

Role of promoter and structural region in autoregulation of nodD

In order to compare the *nodD(1)* genes, including their promoters, of *R. leguminosarum, R. trifolii* and *R. meliloti* with respect to their autoregulatory properties in an isogenic system, the promoter regions and part of the coding regions of these *nodD(1)* genes were cloned in the IncQ transcriptional fusion vector pMP190. This resulted in the plasmids pMP224, pMP192 and pMP195, respectively (Fig. 1A). Each of these plasmids was mobilized into

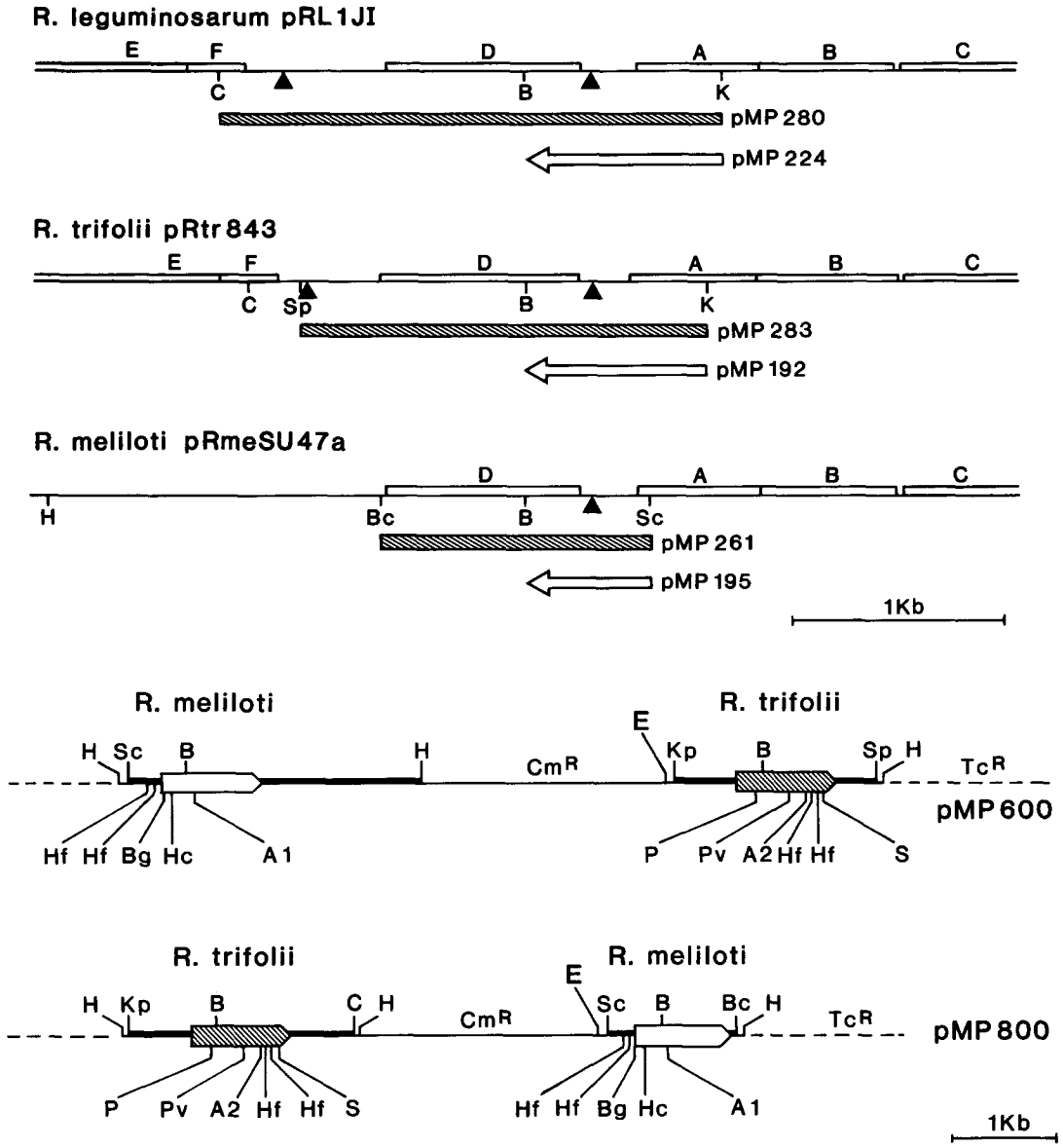

Fig. 1. Construction of plasmids. (A) Inserts containing the *nodD* genes or only their promoters of various *Rhizobium* species are indicated together with the used restriction sites. Size and positions of the indicated *nod* genes are according to published data [31, 30, 8, 34] and data of J. Alan Downie (personal communication) for the translational stop of the *nodA* gene of *R. leguminosarum*. Hatched bars indicate fragments which were cloned in IncP vector pMP92 and arrows indicate the fragments which were cloned in front of *lacZ* in the IncQ transcriptional fusion vector pMP190 [33]. (B) Plasmid pMP600 was constructed by inserting a *Sac* I-*Hind* III fragment derived of pRmSL26 [20] and a *Kpn* I-*Sph* I fragment derived of pRt308 [5], containing the *nodD* genes of *R. meliloti* and *R. trifolii*, respectively, in the vector pMP92 [33], which is a derivative of pTJS75 [29]. In pMP800 the *nodD*-containing fragments were cloned as *Kpn* I-*Cla* I and *Sac* I-*Bcl* I fragments for *R. trifolii* and *R. meliloti*, respectively. The positions of the used restriction sites are indicated in Fig. 1A. Plasmids pMP600 and pMP800 contain the CAT gene from pMP190 [33] cloned as a *Hind* III-*Eco* RI fragment. In the construction of both plasmids several polylinker fragments derived of plasmid pIC20H [22] were used which are not indicated in the figure. Plasmid pMP685 contains a hybrid *nodD* gene which is derived of pMP600 by deletion of the *Bam* HI fragment containing the CAT gene. Indicated are the restriction sites which were used for the characterization of the hybrid *nodD* genes derived from pMP600 and pMP800. Used symbols are A1, *Ava* I; A2, *Ava* II; B, *Bam* HI; Bc *Bcl* I; Bg, *Bgl* II; C, *Cla* I; E, *Eco* RI; H, *Hind* III; Hc, *Hinc* II; Hf, *Hinf* I; K, *Kpn* I; P, *Pst* I; Pv, *Pvu* II; S, *Sma* I; Sc, *Sac* I; Sp, *Sph* I and closed triangle, *nod*-box sequence.

four different *Rhizobium* LPR5045 derivatives, which already harboured one of the *nodD* gene-containing IncP plasmids pMP280, pMP283 and pMP261 (Fig. 1A) or the control plasmid pMP92 in which no *nodD* gene is cloned. This resulted in the construction of a set of 12 strains in which the expression of each of the three *nodD* promoters could be monitored in the absence and presence of each of the three *nodD* genes. From the results (Table 1) it could be concluded that each tested *nodD* product is able to autoregulate and also that each tested *nodD* promoter is susceptible to autoregulation. However, the levels of autoregulation differ strongly, depending on the origin of both the *nodD* gene and the *nodD* promoter. The effectiveness of autoregulation by the *nodD* genes can be arranged in the order: *R. trifolii* > *R. leguminosarum* > *R. meliloti*. This order was shown to be significant using the nonparametric test of Wilcoxon [38] with an α of 5% (data not shown). The *R. meliloti nodDI* product autoregulates only slightly but significantly. The levels of susceptibility of the *nodD* promoters to the *nodD* product are in the same order although in several cases these levels are not significantly different. The autoregulatory properties of *R. trifolii* and *R. leguminosarum* are very similar and are strongly different from those of *R. meliloti* (with *nodDI*) in the homologous combinations.

In vivo *construction of* nodD *hybrid genes and localization of the sites of recombination*

To obtain hybrid *nodD* genes by *in vivo* recombination we have constructed the plasmids pMP600 and pMP800 (Fig.1B) as described in Materials and methods. These plasmids of the IncP class contain the *nodD* gene of *R. trifolii* strain ANU843 and the *nodDI* gene of *R. meliloti* strain 2011 located in tandem, with an unique *Eco* RI restriction site and the gene encoding chloramphenicol-acetyl-transferase (CAT) located in between these genes. Homologous recombination between the *nodD* genes located in tandem will result in plasmids which contain a hybrid *nodD* gene and which no longer possess the *Eco* RI restriction site and the CAT gene. Consequently, *Eco* RI endonuclease treatment of plasmid preparations of pMP600 or pMP800, followed by transformation, enriches for plasmids which contain hybrid *nodD* genes. Using plasmid pMP600, a series of 94 hybrid *nodD* genes, which contain a promoter-terminal moiety of the *R. meliloti nodDI* gene, was obtained and was designated the *mel-tri* series. Using pMP800, in which the positions of the two *nodD* genes is reversed, a series of 40 hybrid *nodD* genes, which contain a 5′ end of the *R. trifolii nodD* gene, was obtained and was designated the *tri-mel* series. By localization of the site of recombination in each hybrid *nodD* gene, using the restriction enzymes indicated in Fig. 1B, the *mel-tri* series of

Table 1. Expression of the *nodD* promoters of *R. leguminosarum*, *R. trifolii* and *R. meliloti* in the presence or absence of *nodD* products of the same sources.

Transcription indicator plasmid with *nodD* promoter of the various sources	β-galactosidase activity ($u \times 10^{-3}$)* in the presence of			
	pMP92 (no *nodD*)	pMP280 (*nodD* of *R. leguminosarum*)	pMP283 (*nodD* of *R. trifolii*)	pMP261 (*nodD* of *R. meliloti*)
pMP244 (*R. leguminosarum*)	3.3 (100%)	1.1 (33%)	0.8 (24%)	2.8 (85%)
pMP192 (*R. trifolii*)	10.4 (100%)	2.8 (28%)	2.1 (20%)	7.8 (75%)
pMP195 (*R. meliloti*)	3.3 (100%)	2.4 (73%)	2.0 (61%)	2.8 (85%)

*Experiments were carried out in quintuplicate. Values in parenthesis are the expression levels relative to the expression in the absence of *nodD* product (pMP92). Standard deviation values of the relative expression values were within 20%.

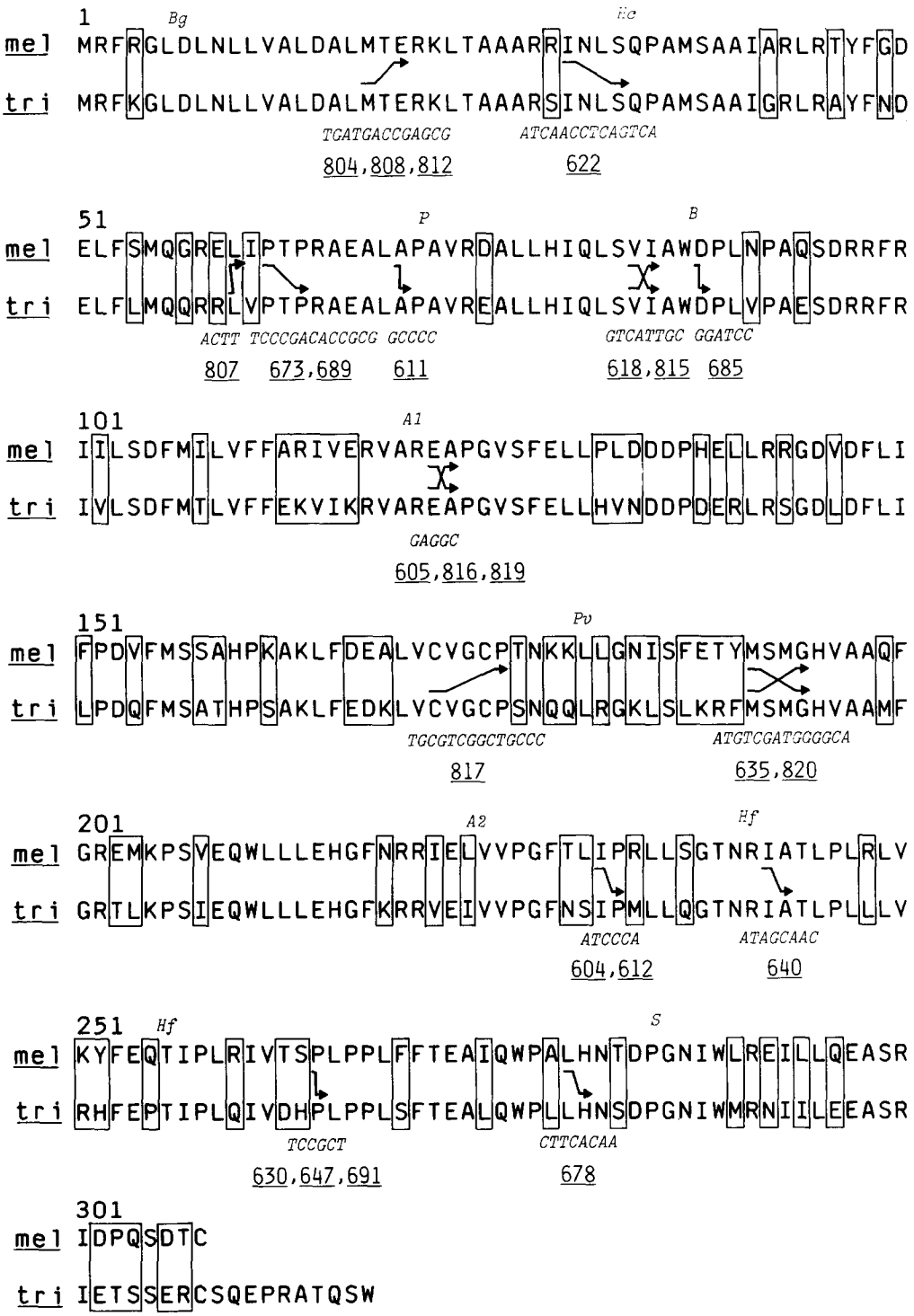

Fig. 2. Localization of the recombination site in *nodD* hybrids. The indicated deduced protein sequences of the *R. meliloti nodD1* and *R. trifolii nodD* products are according to Egelhoff *et al.* [8] and Schofield and Watson [30] respectively. Non-conserved amino acids are given in boxes. Fusion sites of the resulting hybrid *nodD* proteins are indicated by arrows. Arrows down indicate *mel-tri* type of hybrids and arrows up indicate *tri-mel* type of hybrids. The designations of the hybrid genes are indicated under the stretch of homologous nucleotides in which recombination occurred. The length of the arrows indicates the position of the shown stretches of homology. Restriction sites are abbreviated as in Fig. 1.

hybrids could be divided into eight classes and the *tri-mel* series of hybrids into five classes. In addition, the recombination site of at least one representative of each class was localized exactly by nucleotide sequencing. In Fig. 2 the positions of these determined recombination sites are indicated together with the deduced protein sequences of both the *R. trifolii nodD* and *R. meliloti nodD1* genes [30, 8]. The results show that recombination events have occurred in homologous stretches of at least 4 nucleotides long. In none of the sequenced hybrid *nodD* genes were detected any aberrant deletions, duplications or multiple cross-overs. Representative hybrid *nodD* genes were further studied with respect to their autoregulatory properties and their characteristics of activation of the inducible *nod* promoters.

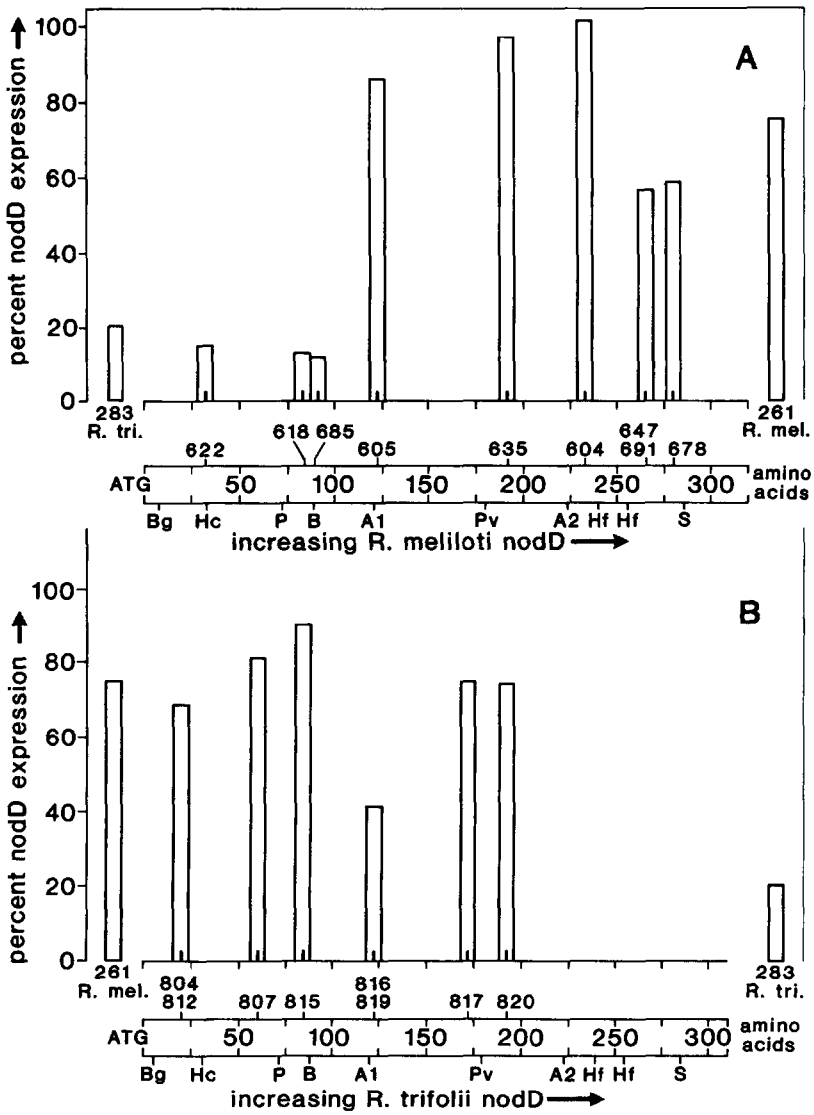

Fig. 3. Autoregulatory properties of various *nodD* hybrids as judged from the expression of the *R. trifolii nodD* promoter. Expression is measured as β-galactosidase activity due to the presence of pMP192. Measured β-galactosidase levels are given relative to the expression level in the absence of a *nodD* product (control plasmid pMP92) and are indicated as a function of the position of the recombination site in the various hybrid *nodD* genes. The horizontal length of the indicated bars is arbitrary. The levels observed with the wild type *nodD* genes are indicated at the two ends of the panels. The position of the recombination sites in the tested hybrids is indicated in the bottom of the figure together with the restriction sites used for the initial classification of the hybrids. For abbreviations, see Fig. 1.

Localization of regions of nodD *involved in autoregulation*

The plasmid pMP192, which contains the *nodD* promoter of *R. trifolii* (Fig. 1A), was used to study the autoregulatory properties of the constructed hybrid *nodD* genes. Therefore representative hybrids of each (restriction) class, were mobilized to strain LPR5045·pMP192 and subsequently the level of transcription of *nodD* in each resulting strain was determined. The results (Fig. 3) show that the *mel-tri* series of hybrids (panel A) can be distinguished into two classes. The first class has autoregulatory properties which are similar to that of the *R. trifolii nodD* gene. The recombination sites of the hybrids of this class are located within the 5′ part of the *nodD* gene which is bordered by the recombination site in *nodD685*. The second class of hybrids autoregulates to a much less extent than the *R. trifolii nodD* product. Hybrids of this class have a recombination site which is located in the 3′ part of *nodD* bordered by the recombination site in *nodD605*. This shows that a region of the *nodD* protein located between amino acids 90 and 118 is involved in autoregulation. The results with the *tri-mel* series (panel B) are less clear. The hybrids *nodD816* and *nodD819* autoregulate significantly stronger than the hybrids with a recombination site which is located more to the 5′ end of *nodD* and this confirms the importance of a region located between amino acids 90 and 118 for autoregulation. However, autoregulation caused by *nodD816* is not identical to that of the *R. trifolii nodD* gene, indicating that also a region located in the carboxy terminal end of the *R. trifolii nodD* product is important for autoregulation.

Ability of hybrid nodD *genes to activate the* R. leguminosarum nodABCIJ *promoter in the presence and absence of luteolin*

The ability of the *mel-tri* and *tri-mel* hybrid genes to activate the inducible *R. leguminosarum nodABCIJ* promoter was tested in the absence and presence of luteolin, a flavonoid which activates both parental *nodD* products efficiently [24, 34]. The expression of the *nodABCIJ* promoter, measured as β-galactosidase activity, is displayed as a function of the position of the recombination site in Fig. 4 (panels A and B). The results obtained with the parental *nodD* genes are given at each end of the panels. Since all other hybrid *nodD* genes within the previously defined recombination classes gave indistinguishable results, only the recombination sites of the indicated representatives of each class were determined exactly by nucleotide sequencing. The results show that the hybrid *nodD* genes are very different from each other in two respects: (i) the level of transcriptional activation in the absence of a flavonoid inducer and (ii) the level of transcriptional activation in the presence of luteolin. Based on these characteristics the hybrids can be divided into four classes. Class I hybrids, represented by the hybrids *nodD635, nodD604* and *nodD640* of the *mel-tri* series, activate the *nodABCIJ* promoter already to the maximally observed level in the absence of an inducer. This level is neither influenced by the presence of luteolin (Fig. 4, Panel A) nor by that of any of the other tested flavonoids (see next section) and therefore this class of *nodD* hybrids will henceforth be referred to as the *nodD* FITA (flavonoid-independent transcription activation) class of genes. The reverse situation, an extremely low level of activation both in the presence and absence of luteolin, was observed with the *tri-mel* series of hybrids represented by *nodD817* and *nodD820*. The recombination site in this class II type of hybrids is located in a similar region of the *nodD* gene (Fig. 4, panel B) as is the case in the FITA class. Class III hybrids behaved similarly to the parental gene products in that activation of the *nodABCIJ* promoter in the absence of luteolin was low whereas the presence of luteolin results in a high level of activation. Surprisingly, class III hybrids of the *tri-mel* series, such as *nodD804*, in which the *nodD* promoter and only one encoded amino acid are different from the *R. meliloti nodD1* gene, are able to exert a substantially higher level of transcriptional activation than this parental *nodD* gene. Class IV hybrids cause a substantial level of transcriptional activation in the absence of inducer and this level is further increased by the presence of luteolin. Although the distinction between the classes III and IV is somewhat arbitrary, it is clear that the most typical examples of class IV

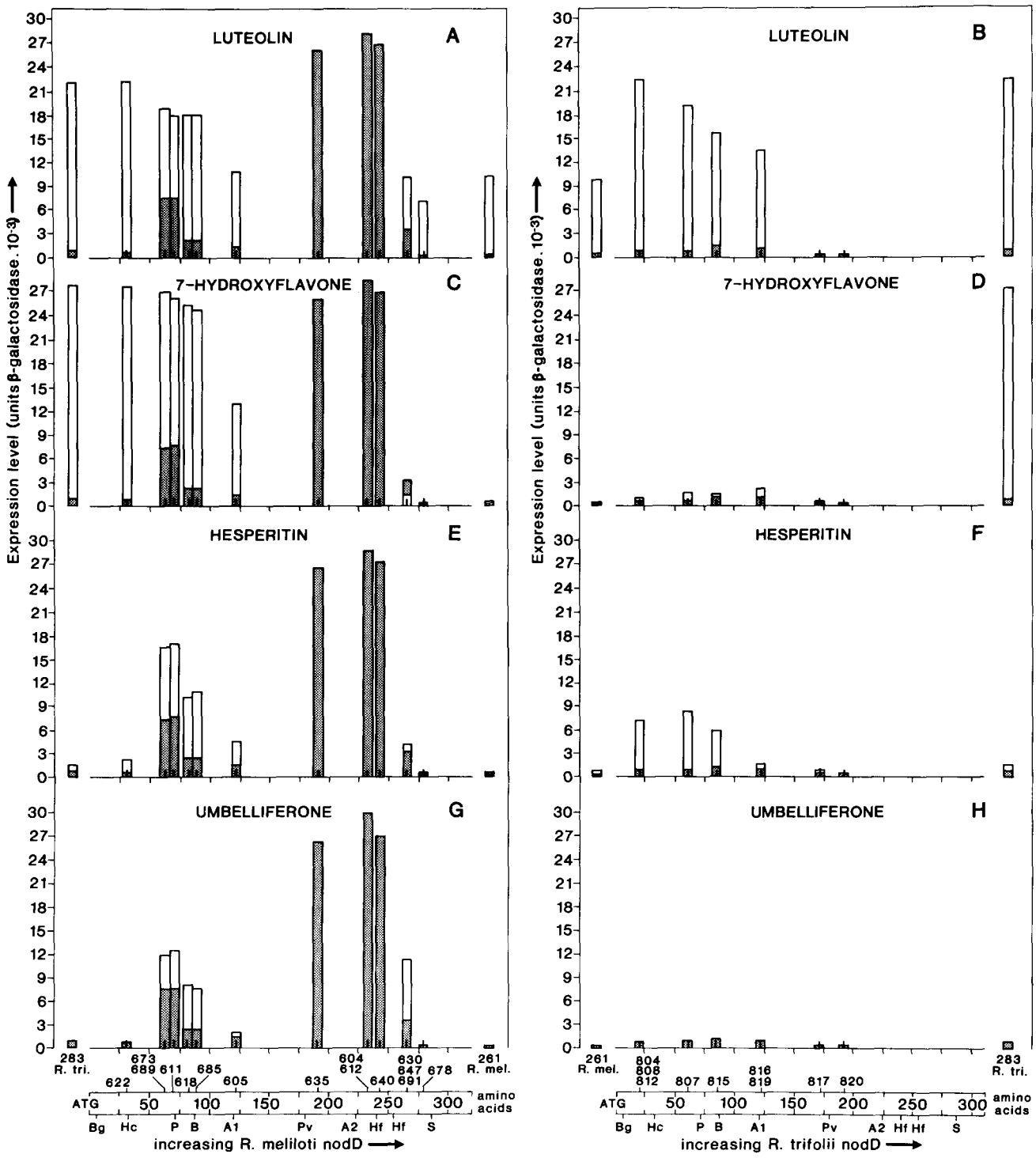

Fig. 4. Expression of the *R. leguminosarum nodABCIJ* promoter in the presence of various *nodD* hybrid genes and in the absence or presence of various flavonoid inducers. The measured expression (in units β-galactosidase $\times 10^{-3}$) with each hybrid *nodD* gene is plotted as a function of the position of the recombination site. The panels A, C, E and G represent the *mel-tri* series of hybrids and the panels B, D, F and H the *tri-mel* series of hybrids. The levels of expression observed with the parental *nodD* genes are indicated at the two ends of the panels. Hatched or open bars, the width of which is chosen arbitrarily, indicate the measured β-galactosidase level in the absence or presence, respectively, of the flavonoids luteolin (panels A and B), 7-hydroxyflavone (panels C and D) and hesperitin (panels E and F) and the coumarin umbelliferone (panels G and H). The position of the recombination site of each hybrid is indicated at the bottom of the figure together with the restriction sites used for the initial classification. For abbreviations, see Fig. 1.

hybrids were found among the *mel-tri* series of hybrids, e.g. *nodD673, nodD611* and *nodD630.* The assignment of the characterized *nodD* hybrids to the defined classes is indicated in Table 2. Since class III and IV hybrids are still efficiently activated by luteolin, a further characterization of these hybrid *nodD* products with respect to their activation by various flavonoids was performed.

Localization of regions involved in the specificity of activation by flavonoids

Activation of the *nodABCIJ* promoter by hybrid *nodD* products was also tested in the presence of the flavone 7-hydroxyflavone and the flavanone naringenin, both of which are able to activate the *nodABCIJ* promoter fully in the presence of the parental *R. trifolii nodD* gene but hardly in the presence of the *R. meliloti nodDI* gene. The results with 7-hydroxyflavone are given in Fig. 4, panels C and D. Similar results were obtained with naringenin (results not shown). The *mel-tri* series of hybrids, with a recombination site in the promoter-terminal part of *nodD* bordered by the recombination site in *nodD605,* can still be activated efficiently by these flavonoids, compared to the induction level in the presence of luteolin (Fig. 4, panel A). However, the *nodD678* product is not activated by 7-hydroxyflavone or naringenin and therefore resembles the *R. meliloti nodDI* product. In the strains harbouring the hybrids *nodD630, nodD647* and *nodD691* 7-hydroxyflavone even inhibits expression. The phenotypic difference between the hybrids

Table 2. Summary of the most relevant properties of the isolated *nodD* hybrids.

Hybrid isolate number	Number of substituted amino acids[1]	Autoregulation type	Activation class[2]	Novel activating substances[3]
mel-tri	compared to *R. trifolii* NodD			
622	2	*R. trifolii*	III	
673, 689	9	not tested	IV	hesp, umb
611	9	not tested	IV	hesp, umb
618	10	*R. trifolii*	IV	hesp, umb
685	10	*R. trifolii*	IV	hesp, umb
605	19	*R. meliloti*	III	hesp
635	44	*R. meliloti*	I	
604, 612	53	*R. meliloti*	I	
640	55	not tested	I	
630, 647, 691	62	*R. meliloti*	IV	umb
678	65	*R. meliloti*	III	
tri-mel	compared to *R. meliloti* NodD			
804, 808, 812	1	*R. meliloti*	III	hesp
807	8	*R. meliloti*	III	hesp
815	10	*R. meliloti*	III	hesp
816, 819	19	intermediate	III	
817	34	*R. meliloti*	II	
820	44	*R. meliloti*	II	

[1] Indicated is the number of amino acids which differ from the parental *nodD* product which constitutes the carboxy terminal part of the hybrid products.

[2] Class I: FITA type; class II: non-functional type; class III: parental type; class IV: substantial level of transcriptional activation in the absence of inducer.

[3] Abbreviations: hesp, hesperitin; umb, umbelliferone.

nodD605 and *nodD678* indicates that a region located between amino acids 132 and 270 is involved in the specificity of activation by 7-hydroxyflavone and naringenin. This conclusion is supported by the results with the *tri-mel* series of hybrids since the hybrid *nodD816* and *nodD819* products are still not efficiently activated by 7-hydroxyflavone and naringenin. Since inducers specific for the *R. meliloti nodD1* product are not known, this approach could not be used to test this conclusion. A more precise localization of this specificity region is obstructed by the occurrence of the *nodD* FITA phenotype in the *mel-tri* series and of the non-functional *nodD* phenotypes in the *tri-mel* series.

Since several types of hybrids had surprising novel properties we also tested whether hybrids are activated by flavonoids which are unable to activate either of the two parental *nodD* products. Therefore the hybrid *nodD* genes were tested in the presence of the flavanone hesperitin, the isoflavones daidzein and genistein, and the coumarin umbelliferone, neither of which can activate the inducible *nod* promoter to a substantial level in the presence of the parental *nodD* proteins. Moreover daidzein, genistein and umbelliferone have been shown to even inhibit the induction by flavonoids [9, 6]. As shown in Fig. 4, panels E and F, several hybrid products, e.g. of *nodD673, nodD605, nodD804* and *nodD815*, appeared to be activated by the flavanone hesperitin. In the *mel-tri* series (panel E) this concerns the hybrid *nodD* products with the fusion site located between amino acids 41 and 122. In the *tri-mel* series (panel F) this concerns the hybrid *nodD* products with the fusion site located between amino acids 4 and 75. We therefore must conclude that also several regions in the N-terminal part of the *nodD* product, bordered by amino acid 122, are involved in the determination of the specificity of activation by flavonoids. It appeared that transcriptional activation by the hybrid *nodD* products was not affected, positively or negatively, by the addition of the isoflavones daidzein or genistein (results not shown). However, as shown in Fig. 4, panel G, several hybrids of the *mel-tri* series appeared to be activated by the coumarin umbelliferone. In contrast, none of the hybrids of the *tri-mel* series was activated with umbelliferone (panel H). The activation of the hybrid products of *nodD673, nodD689, nodD611, nodD618* and *nodD685* with umbelliferone confirm the importance of the amino-terminal region of the *nodD* product in the determination of flavonoid specificity.

Discussion

The interaction between flavonoid inducer and *nodD* product is crucial in the nodulation process. Recently, it has been shown that the *nodD* product is able to bind to DNA sequences upstream of the flavonoid-inducible *nod* operons [12, 10, E. Kondorosi, personal communication]. Furthermore, it has been shown that the origin of the *nodD* product in isogenic *Rhizobium* or *Agrobacterium* strains determines which flavonoids act as inducers [34, 41]. These data support a model in which flavonoids bind to the *nodD* product which can result in the conversion to an activated form which then stimulates transcription of the inducible *nod* operons. As proposed by Djordjevic *et al.* [6] non-inducing compounds like isoflavones or coumarins might also bind to the *nodD* product but are unable to convert it to an activated form in this way inhibiting induction by other flavonoids. In order to study the various properties of *nodD* in more detail two series of hybrid *nodD* genes have been constructed, derived of the *nodD(1)* genes of *R. trifolii* and *R. meliloti*. This approach was used to localize regions of the *nodD* product involved in the functional differences between these *nodD* genes, namely in their ability to "recognize" a particular flavonoid as an inducer, to activate the inducible *nod* promoters and to autoregulate their own constitutive transcription. Besides a localization of several regions involved in these differences, also various *nodD* hybrid genes were obtained whose characteristics of transcriptional activation were strongly different from the parental *nodD* genes. These *nodD* genes with novel properties might be useful for studying the mechanism of action of the *nodD* product in future research. The results show that the construction of hybrid *nodD* genes can be used to scan regions in the entire *nodD* protein for their involvement in its functions. However, as this method only can study the

role of the non-conserved amino acids in the parental genes it must be considered as complementary to random mutagenesis.

Both the promoter and the structural region of the *nodD* gene play a role in determining the differences in the level of autoregulation observed between the *nodD(1)* genes of *R. leguminosarum, R. trifolii* and *R. meliloti* (Table 1). The results of Fig. 3 indicate that the 10 non-conserved amino acids, located in the amino-terminal part of the *nodD* protein up to amino acid 90, are not involved in the difference in autoregulation between the *nodD* protein of *R. trifolii* and the *nodD1* protein of *R. meliloti.* Two regions of the *nodD* product, separated by a stretch of conserved amino acids, were identified which are involved in determining the level of autoregulation: a region located between amino acids 90 and 118 and a region located between amino acid 131 and the carboxy-terminal end of the protein. The first region contains only 9 amino acids which differ in the parental *nodD* products (Fig. 2), which makes this region attractive to study in more detail using site-directed mutagenesis. The significance of the second region is not clear. Its identification is mainly based on the results obtained with the *tri-mel* series of hybrids, of which only few hybrids were isolated which resulted from a recombination in this region. The low level of autoregulation of the hybrids *nodD817* and *nodD820* suggests that the second region is essential for autoregulation. However, it cannot be excluded that the absence of autoregulation observed with these hybrids is not directly linked with the process of autoregulation but is an indication that the products of these hybrid genes are completely non-functional since they are also unable to activate the inducible *nodABCIJ* promoter (see below).

The hybrid *nodD* genes were tested for their ability to activate the inducible *nodABCIJ* promoter of *R. leguminosarum* in the presence and in the absence of the flavone luteolin, a common activator of both parental *nodD* products. Two classes of *nodD* hybrids were identified whose characteristics were no longer comparable with those of the wild-type *nodD* genes. Representatives of the first class activate the inducible *nod* promoter to a maximum level in the absence of inducer. This level could not be influenced, positively or negatively, by the addition of known inducers or anti-inducers, and therefore these *nodD* hybrids were designated as FITA (flavonoid independent transcription activation) genes. Representatives of the second class of *nodD* hybrids cannot activate the inducible *nod* promoter with any of the tested compounds and should therefore be considered as non-functional *nodD* products. The experiments described in this paper have been carried out using the *R. leguminosarum nodABCIJ* promoter as a probe for transcriptional activation. The characteristics of the FITA and non-functional hybrids were also observed when the inducible promoters of the *nodABC(IJ)* genes of *R. trifolii* or *R. meliloti* were used as transcriptional indicators (data not shown). The significance of the FITA phenotype for the nodulation process is presently under investigation in our laboratory. The determined recombination sites of the *mel-tri* series of hybrids indicate that the FITA characteristic is the result of a combination of a region originating from the *R. meliloti nodD1* product, located between amino acid 132 and 191, and a region originating from the *R. trifolii nodD* product, located between amino acids 248 and 264. Although these two regions are separated by a stretch of 57 amino acids in the primary structure, they may be part of the same functional domain as a result of folding of the *nodD* product. The localized sites of recombination in the *tri-mel* series hybrids indicate that the non-functional *nodD* phenotype is caused by the combination of a region originating from the *R. trifolii nodD* product, which is located between amino acids 132 and 169, and a region located within the *R. meliloti nodD1* terminal part of *nodD820.* Although the right border of the latter region is not defined by a recombination site, one can see a resemblance with the regions which cause the *nodD* FITA phenotype, which suggests that both phenotypes are the result of changes affecting the same functional domain.

The results (Fig.4) indicate that a region located in the carboxy-terminal moiety of the *nodD* product is involved in the specificity of activation by flavonoids. The results obtained with both series of hybrid genes, using the *R. trifolii nodD* activators 7-hydroxyflavone and naringenin, localize this region between amino acids 132 and 270. A more precise

localization was not possible because of the FITA phenotype and the non-functional phenotype of hybrids resulting from recombination in this region. This identified region is consistent with, and extends, the observation of Horvath *et al.* [17] who indicated a region located between the carboxy-terminal end and amino acid 90 to be a determinant of exudate specificity. Burn *et al.* [1] have described a point mutation of the *R. leguminosarum nodD* product which shows a changed flavonoid specificity and the ability to activate the *nod* promoter to a substantial level in the absence of flavonoid inducer. This mutation is a substitution of an aspartic acid at position 284 of the wild-type gene for asparagine. Since our method can only show the importance of amino acids which differ between *nodD* proteins, and since amino acid 284 is conserved between the *nodD* proteins we chose to use, we can neither confirm nor contradict the importance of this amino acid. It is striking that several of our hybrids, such as *nodD673, nodD611* and *nodD630,* resemble the phenotype of the mentioned point mutation with respect to the transcriptional activation in the absence of inducer and the ability to be activated by substances which are known as anti-inducers (Fig. 4, panels A and G).

The presumption that the specificity of exudate recognition is only determined by the carboxy-terminal region of *nodD* is consistent with the high conservation of the amino-terminal region of the *nodD* product [17]. In strong contrast, our results with the flavanone hesperitin and the coumarin umbelliferone, as summarized in Table 2, indicate that both the carboxy-terminal region and regions located in the amino-terminal part, left of amino acid 90, are involved in this specificity. Very surprisingly, replacement of amino acid arginine at position four of the *R. meliloti nodDI* product by the functionally similar amino acid lysine alters its flavonoid specificity. Hybrids which contain this substitution also have the original *nodD* promoter substituted for that of *R. trifolii.* However, it is unlikely that the changed specificity for flavonoids of the *nod804* product is due to a promoter substitution, since replacement of the *R. trifolii* promoter of *nodD804* by other promoters resulted in *Rhizobium* strains with the same flavonoid specificity (results not shown). Other regions, located in the amino-terminal moiety of the *nodD* product, which are involved in flavonoid specificity are less clearly defined. By comparing the phenotypes of *nodD* hybrids of which the products differ only few amino acids, we can conclude that at least two other segments of this moiety, located between amino acids 41 and 62 (as deduced from the results with the *mel-tri* series) and between amino acids 90 and 118 (as deduced from the results with the *tri-mel* series), are involved in this specificity. In conclusion, our results show that there are many regions in the *nodD* product which are involved in the specificity of activation by flavonoids. Furthermore, the regions determined to be involved in autoregulation are not clearly separated from regions involved in the flavonoid specificity.

Acknowledgements

We thank Marcel Hoffmann, Henk Roest and Jan Verdoes for their help in part of the experiments and Robert Fisher for critically reading the manuscript.

References

1. Burn J, Rossen L, Johnston AWB: Four classes of mutations in the *nodD* gene of *Rhizobium leguminosarum* biovar. *viciae* that affect its ability to autoregulate and/or activate other nod genes in the presence of flavonoid inducers. Genes & Development 1: 456–464 (1987).
2. Debellé F, Sharma SB: Nucleotide sequence of *Rhizobium meliloti* RCR2011 genes involved in host specificity of nodulation. Nucl Acids Res 14: 7453–7472 (1986).
3. Ditta G, Stanfield S, Corbin D, Helinski DR: Broad host-range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti.* Proc Natl Acad Sci USA 77: 7347–7351 (1980).
4. Djordjevic MA, Schofield PR, Rolfe BG: Tn5 mutagenesis of *R. trifolii* host-specific nodulation genes result in mutants with altered host-range ability. Mol Gen Genet 200: 463–471 (1985).
5. Djordjevic MA, Innes RW, Wijffelman CA, Schofield PR, Rolfe BG: Nodulation of specific legumes is controlled by several distinct loci in *Rhizobium trifolii.* Plant Mol Biol 6: 389–401 (1986).
6. Djordjevic MA, Redmond JW, Batley M, Rolfe BG: Clovers secrete specific phenolic compounds which either stimulate or repress *nod* gene expression in *Rhizobium trifolii.* EMBO J 6: 1173–1179 (1987).

7. Djordjevic MA, Gabriel DW, Rolfe BG: *Rhizobium* – the refined parasite of legumes. Ann Rev Phytopath 25: 145–168 (1987).
8. Egelhoff TT, Fisher RF, Jacobs TW, Mulligan JT, Long SR: Nucleotide sequence of *Rhizobium meliloti* 1021 nodulation genes: *nodD* is read divergently from *nodABC*. DNA 4: 241–248 (1985).
9. Firmin JL, Wilson KE, Rossen L, Johnston AWB: Flavonoid activation of nodulation genes in *Rhizobium* reversed by other compounds present in plants. Nature (London) 324: 90–92 (1986).
10. Fisher RF, Egelhoff T, Mulligan JT, Long SR: Specific binding of proteins from *Rhizobium meliloti* cell-free extracts containing nodD to DNA sequences upstream of inducible nodulation genes. Genes & Development 2: 282–293 (1988).
11. Göttfert M, Horvath B, Kondorosi E, Putnoky P, Rodriguez-Quinones F, Kondorosi A: At least two different *nodD* genes are necessary for efficient nodulation on alfalfa by *Rhizobium meliloti*. J Mol Biol 191: 411–420 (1986).
12. Hong GF, Burn JE, Johnston AWB: Evidence that DNA involved in the expression of nodulation (*nod*) genes in *Rhizobium* binds to the product of the regulatory gene *nodD*. Nucl Acids Res 15: 9677–9690 (1987).
13. Honma MA, Ausubel FM: *Rhizobium meliloti* has three functional copies of the *nodD* symbiotic regulatory protein. Proc Natl Acad Sci USA 84: 8558–8562 (1987).
14. Hooykaas PJJ, Klapwijk PM, Nuti MP, Schilperoort RA, Rörsch A: Transfer of the *Agrobacterium* Ti plasmid to avirulent agrobacteria and to rhizobia ex planta. J Gen Microbiol 98: 477–484 (1977).
15. Hooykaas PJJ, Schnijdewindt FGM, Schilperoort RA: Identification of the Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its transfer to and expression in other Rhizobia and *Agrobacterium tumefaciens*. Plasmid 8: 73–82 (1982).
16. Horvath B, Kondorosi E, John M, Schmidt J, Török I, Györgypal Z, Barabas I, Wieneke U, Schell J, Kondorosi A: Organisation, structure and symbiotic function of *Rhizobium meliloti* nodulation genes determining host specificity for alfalfa. Cell 46: 335–343 (1986).
17. Horvath B, Bachem CW, Schell J, Kondorosi A: Host-specific regulation of nodulation genes in *Rhizobium* is mediated by a plant-signal, interacting with the *nodD* gene product. EMBO J 6: 841–848 (1987).
18. Innes RW, Kuempel PL, Plazinski J, Canter Cremers HCJ, Rolfe BG, Djordjevic MA: Plant factors induce expression of nodulation and host-range genes in *R. trifolii*. Mol Gen Genet 201: 426–432 (1985).
19. Kieny MP, Lathe R, Lecocq JP: New versatile cloning and sequencing vectors based on bacteriophage M13. Gene 26: 91–99 (1983).
20. Long SR, Buikema WJ, Ausubel FM: Cloning of *Rhizobium meliloti* nodulation genes by direct complementation of Nod-mutants. Nature (London) 298: 485–488 (1982).
21. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring, NY (1982).
22. Marsh JL, Erfle M, Wykes EJ: The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene 32: 481–485 (1984).
23. Mulligan JT, Long SR: Induction of *Rhizobium meliloti nodC* expression by plant root exudate requires *nodD*. Proc Natl Acad Sci USA 82: 6609–6613 (1985).
24. Peters NK, Frost JW, Long SR: A plant flavone, luteolin, induces expression of *Rhizobium meliloti* genes. Science 233: 977–980 (1986).
25. Redmond JW, Batley M, Djordjevic MA, Innes RW, Kuempel PL, Rolfe BG: Flavones induce expression of nodulation genes in *Rhizobium*. Nature (London) 323: 632–635 (1986).
26. Rossen L, Shearman CA, Johnston AWB, Downie JA: The *nodD* gene of *Rhizobium leguminosarum* is autoregulatory and in the presence of plant exudate induces the *nodA,B,C* genes. EMBO J 4: 3369–3373 (1985).
27. Rostas K, Kondorosi E, Horvath B, Simoncsits A, Kondorosi A: Conservation of extended promoter regions of nodulation genes in *Rhizobium*. Proc Natl Acad Sci USA 83: 1757–1761 (1986).
28. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
29. Schmidhauser TJ, Helinski DR: Regions of broad-host-range plasmid RK2 involved in replication and stable maintenance in nine species of gram-negative bacteria. J Bact 164: 446–455 (1985).
30. Schofield PR, Watson JM: DNA sequence of *Rhizobium trifolii* nodulation genes reveals a reiterated and potentially regulatory sequence preceding *nodABC* and *nodFE*. Nucl Acids Res 14: 2891–2903 (1986).
31. Shearman CA, Rossen L, Johnston AWB, Downie JA: The *Rhizobium leguminosarum* nodulation gene *nodF* encodes a polypeptide similar to acyl-carrier protein and is regulated by *nodD* plus a factor in pea root exudate. EMBO J 5: 647–652 (1986).
32. Spaink HP, Okker RJH, Wijffelman CA, Pees E, Lugtenberg B: Promoters and operon structure of the nodulation region of the *Rhizobium leguminosarum* symbiosis plasmid pRL1JI. In: Lugtenberg B (ed.) Recognition in Microbe-Plant Symbiotic and Pathogenic Interactions, NATO ASI Series, Vol H 4, pp. 55–68. Springer-Verlag, Berlin/Heidelberg (1986).
33. Spaink HP, Okker RJH, Wijffelman CA, Pees E, Lugtenberg BJJ: Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. Plant Mol Biol 9: 27–39 (1987).
34. Spaink HP, Wijffelman CA, Pees E, Okker RJH, Lugtenberg BJJ: *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. Nature (London) 328: 337–340 (1987).
35. Swanson JA, Tu JK, Ogawa J, Sanga R, Fisher RF, Long SR: Extended region of nodulation genes in *Rhizobium meliloti* 1021. I. Phenotype of Tn5 insertion mutants. Genetics 117: 181–189 (1987).
36. Tommassen J, Van der Ley P, Van Zeijl M, Achterberg M: Localization of functional domains in *E. coli* K12 outer membrane porins. EMBO J 4: 1583–1587 (1985).

37. Wijffelman C, Zaat B, Spaink H, Mulders I, Van Brussel A, Okker R, De Maagd R, Lugtenberg B: Induction of *Rhizobium nod* genes by flavonoids: differential adaptation of promoter, *nodD* gene and inducers for various cross-inoculation groups. In: Lugtenberg B (ed) Recognition in Microbe-Plant Symbiotic and Pathogenic Interactions, NATO ASI Series, Vol H 4, pp. 123–135. Springer-Verlag, Berlin/Heidelberg (1986).
38. Wilcoxon F: Probability tables for individual comparisons by ranking methods. Biometrics 3: 119–122 (1947).
39. Yanisch-Perron C, Vieira J, Messing J: Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC vectors. Gene 33: 103–119 (1985).
40. Zaat SAJ, Wijffelman CA, Spaink HP, van Brussel AAN, Okker RJH, Lugtenberg BJJ: Induction of the *nodA* promoter of the *R. leguminosarum* Sym plasmid pRL1JI by plant flavonones and flavones. J Bact 169: 198–204 (1987).
41. Zaat SAJ, Wijffelman CA, Mulders IHM, van Brussel AAN, Lugtenberg BJJ: Root exudates of various host plants of *R. leguminosarum* contain different sets of inducers of *Rhizobium* nodulation genes. Plant Physiol 86: 1298–1303 (1988).