

MicroCorrespondence

Structure of the DNA-binding domain of the OmpR family of response regulators

Sir,

The *Escherichia coli* DNA-binding protein, OmpR, is one of the best characterized of the bacterial positive regulators that enhance the transcriptional ability of RNA polymerase. Furthermore, OmpR is a typical example of response regulators involved in the widespread two-component phosphotransfer signalling systems (Parkinson and Kofoed, 1992, *Annu Rev Genet* 26: 71–112). OmpR contains a phospho-accepting receiver domain of approx. 120 amino acids at the N-terminal portion, followed by a C-terminal DNA-binding domain of approx. 120 amino acids. An inspection of current databases (SWISSPROT/PIR/GenBank) revealed that *E. coli* has at least 15 members of the OmpR family of proteins, including one encoded by an *E. coli* plasmid, whose amino acid sequences show extensive similarities to each other in their C-terminal DNA-binding domains as well as the N-terminal receiver domains. The occurrence of members of the OmpR family has also been reported frequently for many other bacterial species. These facts suggest that the OmpR-like proteins are one of the most widespread DNA-binding transcriptional regulators in bacteria.

The members of the OmpR family in *E. coli* and related species have long been the subject of debate in this journal (OmpR, Mizuno and Mizushima, 1990, *Mol Microbiol* 4: 1077–1082; KdpE, Nakashima *et al.*, 1993, *Mol Microbiol* 7: 109–116; PhoP, Wick *et al.*, 1995, *Mol Microbiol* 16: 465–47; ArcA, Drapal and Sawers, 1995, *Mol Microbiol* 16: 597–607; TorR, Simon *et al.*, 1995, *Mol Microbiol* 17: 971–980). One of the crucial issues concerning the OmpR family of proteins is the molecular mechanism underlying DNA recognition. A previous inspection of the primary structure of OmpR failed to find a canonical helix-turn-helix (HTH) DNA-binding motif, which is commonly found in a number of *E. coli* DNA-binding proteins. Furthermore, these OmpR-like proteins are highly similar to each other in their presumed DNA-binding domains, yet they are postulated to regulate (activate or repress) specific subsets of genes in *E. coli*. The other interesting issue is the underlying molecular mechanism of gene activation *per se*. It has been suggested that the DNA-binding domain of OmpR contains a direct contact site(s) with RNA polymerase (particularly with the α -subunit). The three-dimensional structure of OmpR, which would help us

study these relevant issues in more depth, is a long time in coming. The structure of the C-terminal DNA-binding domain of OmpR, determined by X-ray crystallography, is now available (Kondo *et al.*, 1997, *Nature Struct Biol* 4: 28–31; Martinez-Hackert and Stock, 1997, *Structure* 5: 109–124). The determined secondary structure of the C-terminal domain of OmpR (residues 137–234) is shown schematically in Fig. 1. This structural view also allows us to unambiguously align the sequences of all the *E. coli* OmpR family, and should be useful, particularly for studying the OmpR family of *E. coli* and other species (note that the additional member, CadC, is atypical in the sense that it does not belong to the family of response regulators).

With special reference to the DNA recognition by the OmpR family of proteins, the X-ray structure reveals that two helices, $\alpha 2$ and $\alpha 3$, create a structure that superimposes well on those of other canonical HTH motifs (e.g. CRP, cAMP receptor protein). However, this presumed HTH structure of OmpR is unique in the sense that it has a long loop between $\alpha 2$ and $\alpha 3$. This may be the reason why previous surveys failed to detect a putative HTH motif in OmpR. One can now envisage that $\alpha 3$ of OmpR is the DNA-recognition helix, as is the case for the standard HTH-type DNA-binding proteins.

This presumed recognition helix, $\alpha 3$, consisting of 12 amino acid residues, is enlarged in Fig. 2, along with those of other members of the OmpR family of proteins. It is clear that hydrophobic residues (hydrophobic a.a.) at positions 1, 5, 8 and 12 are well conserved among the OmpR protein family (Fig. 2, left-hand panel). When these residues are depicted on an α -helical wheel, they create a hydrophobic face along one side of the helical cylinder (Fig. 2, right-hand panel). In the determined structure of OmpR, this hydrophobic side is facing the protein core. If $\alpha 3$ is the DNA-recognition helix, as presumed, other residues in $\alpha 3$ may be facing DNA (or bases in the major groove). Residues at positions 7, 9 and 11 are frequently positively charged amino acids (+ a.a.). As indicated in Fig. 2, these residues on the equator may be involved in an interaction with the negatively charged DNA backbones. The remaining residues, at positions 3, 6 and 10, may be in close proximity to specific bases, thereby enabling OmpR to recognize a specific target DNA. The DNA-binding specificity for each member may be determined, at least in part, by these amino acids (note that the amino acid combination in OmpR, i.e. Val-3, Ser-6 and Arg-10, is not identical to the amino acid combination

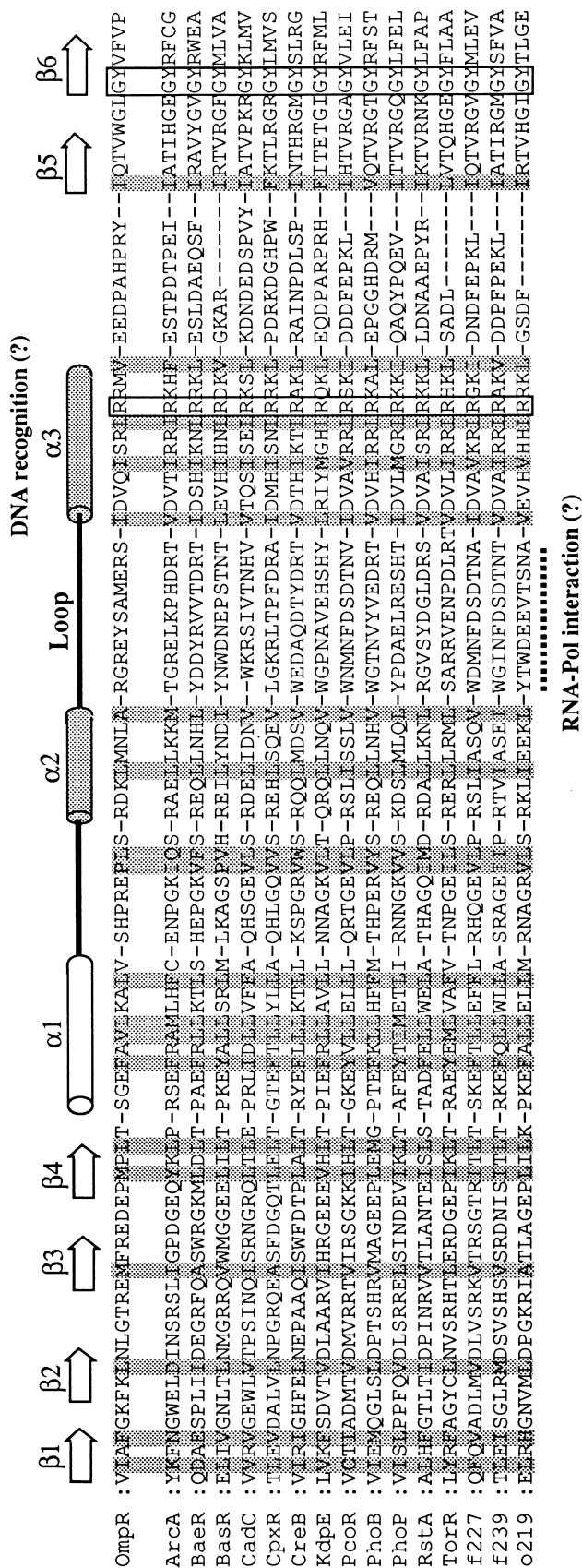


Fig. 1. Alignment of the amino acid sequences of the *E. coli* OmpR family of proteins on the basis of the structure of the C-terminal DNA-binding domain of OmpR. The C-terminal OmpR sequence, extending from positions 137–234, is shown. Above this sequence, the secondary structure of OmpR, determined by X-ray crystallography, is depicted schematically (i.e. three α -helices are denoted by $\alpha 1$ to $\alpha 3$, and six β -strands are denoted by $\beta 1$ to $\beta 6$). Hydrophobic residues conserved among the aligned members are shaded. Three residues that are invariant among the members are boxed. Some spaces are introduced into the aligned sequences in order to emphasize the structural features. The sequences are from the following databases (accession number): OmpR, SWISSPROT (SP) (P03025); ArcA, SP (P30846); Baer, SP (P03026); Basr, SP (P30846); Cadc, SP (P23890); CpxR, SP (P16244); Creb, SP (P08368); KdpE, GenBank (GB) (ECAA000173); PcoR, PIR (S52257); PhoB, SP (P08402); PhoP, SP (P23836); RstA, GB (U41101); TorR, SP (P38684); f227, GB (ECAA000162); f239, GB (ECAA000288); o219, GB (ECAA000384).

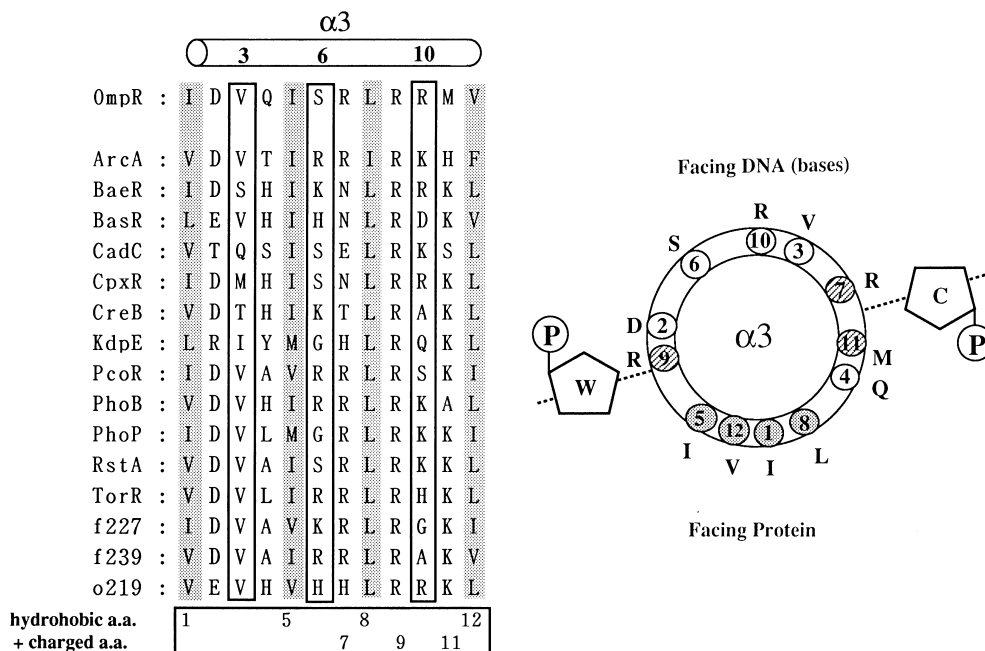
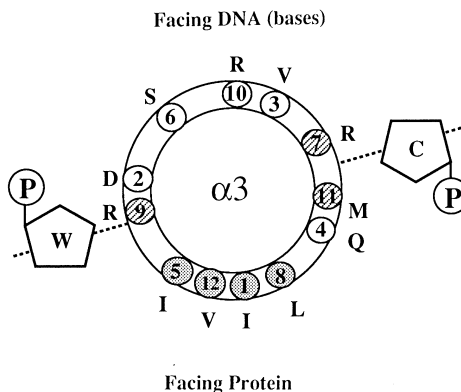


Fig. 2. The putative DNA-recognition helix in the C-terminal DNA-binding domain of OmpR. The $\alpha 3$ sequence of the putative DNA-recognition helix of OmpR is shown, along with those of other members of the OmpR family. The $\alpha 3$ sequence, consisting of 12 amino acid residues, is also plotted in an α -helical wheel. This helix is supposed to interact with the major groove of DNA. (DNA backbones are indicated by pentagons; W, Watson strand; C, Crick strand.) Other details are given in the text.

found at these positions in any other members). It may also be worth mentioning that a classical OmpR mutant, OmpR2, which exhibits an OmpF-constitutive and OmpC-negative phenotype, turned out to have an amino acid substitution at position 3 (Val-203→Met).

Regarding the contact domain with RNA polymerase, a set of positive control (PC) mutants have previously been reported for OmpR (Pratt and Silhavy, 1994, *J Mol Biol* **243**: 579–594; Koto *et al.*, 1995, *Mol Gen Genet* **248**: 399–406). On the basis of the determined structure, the amino acid substitutions in these PC-mutants are located in the loops between $\alpha 2$ and $\alpha 3$ (Glu-193, Ala-196, Glu-198), or $\alpha 1$ and $\alpha 2$ (Pro-179, Ser-181) (see Fig. 1). Such loop regions in the OmpR family of proteins are thus considered to be involved in an interaction with RNA polymerase (either with the α - or σ -subunit).

In conclusion, a close inspection of these amino acid



sequences of OmpR-like proteins, in the context of the determined OmpR structure, should provide us with new and at least preliminary insights into the structure and function of the members of the large OmpR family of proteins.

Takeshi Mizuno^{1*} and Isao Tanaka²

¹Laboratory of Molecular Microbiology, School of Agriculture, Nagoya University, Chikusa-ku, Nagoya 464, Japan.

²Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060, Japan.

*For correspondence.

E-mail i45455a@nucc.cc.nagoya-u.ac.jp; Tel. (52) 789 4089. Fax (52) 789 4091.

Received 5 December, 1996; revised 18 February, 1997; accepted 26 February, 1997

***Neisseria gonorrhoeae* reverse transcriptase activity does not mediate pilin gene conversion**

Sir,

Pilin variation in *Neisseria gonorrhoeae* is implicated in both immune evasion and modulation of the bacterial adherence to human host cells. Frequent changes in the

main subunit of the pili are caused by homologous recombination events between the expression locus *pilE* and several variant silent pilin genes (*pilS*) lacking a promoter and 5'-terminal codons. Pilin variation can occur by either transformation or intra-strain recombination (for a review see Robertson and Meyer, 1992, *Trends Genet* **8**: 422–427; Seifert, 1996, *Mol Microbiol* **21**: 433–440). Intriguingly, the latter pathway usually results in non-reciprocal,

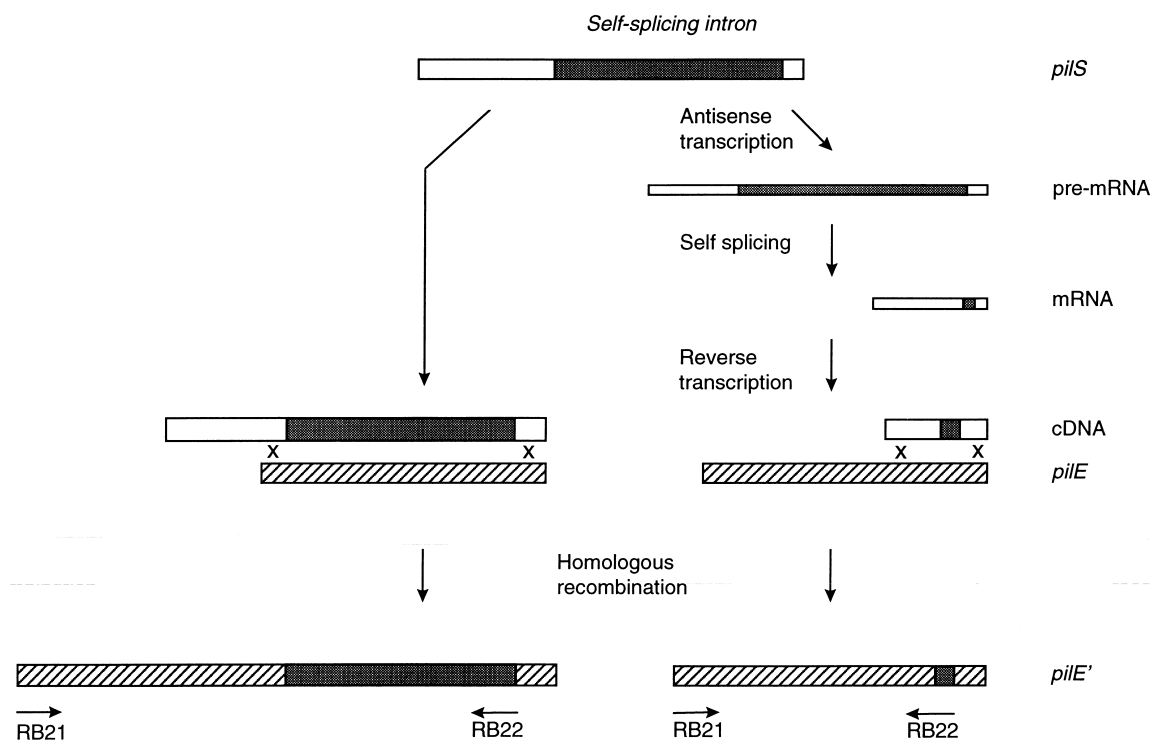


Fig. 1. Hypothetical pathways that could account for non-reciprocal recombination (gene conversion) with *pilE1*. The donor locus (*pilS*) was modified so that it contains a self-splicing intron. If gene conversion involves the formation of a precursor transcript (pre-mRNA) of *pilS*, self-splicing should occur and recombination of the resulting cDNA should result in a small insertion in *pilE* (right-hand panel). If no RNA intermediate exists, gene conversion should lead to a large insertion in *pilE* (left-hand panel). Using the primer pair RB21 (5'-GGCTTCCCCTTCAATTAGGAG-3') and RB22 (5'-TCCGAATCCCACCTCT-3') for PCR, specific gene-conversion events between the *pilS1* and *pilE1* could be readily detected and the outlined recombination pathways distinguished.

gene-conversion-like changes. In yeast cells, gene-conversion events can be mediated by the synthesis of a diffusible reverse transcript and its subsequent recombination into the chromosome (Derr and Strathern, 1993, *Nature* **361**: 170–173). In addition, prokaryotic reverse transcriptases (RTs) are able to synthesize double-stranded DNA (Lampson *et al.*, 1990, *J Biol Chem* **265**: 8490–8496), and homologous recombination can be promoted by RT activity *in vitro* (Negróni *et al.*, 1995, *Proc Natl Acad Sci USA* **92**: 6971–6975). We thus addressed the questions of whether gonococci possess detectable RT activity and if the observed gene conversions are mediated by reverse transcription.

The soluble fraction obtained from bacterial lysates (Brutlag *et al.*, 1971, *Proc Natl Acad Sci USA* **68**: 2826–2829) was assessed for RT activity using the highly specific assay described by Silver *et al.* (1993, *Nucleic Acids Res* **21**: 3593–3594). Single-stranded Brome Mosaic Virus (BMV) RNA (Promega) was reverse transcribed to yield cDNA using a specific primer, and then amplified via a polymerase chain reaction (PCR) using two other primers derived from the BMV sequence. Boiled cell extracts served as a negative control, whereas avian myeloblastosis virus (AMV) RT (USB) was used as positive control

for the reverse transcription of the RNA into DNA. Both untreated gonococcal extracts and purified *Escherichia coli* DNA polymerase I (Pharmacia) yielded the predicted PCR product which is indicative of RT activity (data not shown). Whether this measured RT activity is mediated by a gonococcal DNA polymerase I or by a distinct RT is, at present, unknown. Multicopy single-stranded DNA (msDNA) associated with prokaryotic reverse transcriptase activity has previously not been detected in gonococci (Innouye and Innouye, 1991, *Annu Rev Microbiol* **45**: 163–186).

In order to test the hypothesis that gene conversion in gonococci is based on the reverse transcription of *pilS* genes, we developed a genetic system to monitor such events. N969, a derivative of *N. gonorrhoeae* strain MS11, was mutated in *comA*, leading to transformation deficiency so that transformation-mediated pilin variation was excluded (Faciús and Meyer, 1993, *Mol Microbiol* **10**: 699–712). A cloned invariant *pilC* gene (Rudel *et al.*, 1995, *Nature* **373**: 357–359) ensured that switching between pilated (P^+) and non-piliated (P^-) phenotypes was not due to phase variation of PilC expression. Furthermore, a stop signal, an RB22-primer-binding site, and a self-splicing intron from *Tetrahymena* (generously provided by R. B. Waring

(Waring, 1989, *Nucleic Acids Res* **17**: 10281–10293) were cloned into the hypervariable region (mc2) (Haas *et al.*, 1992, *Mol Microbiol* **6**: 197–208) of *pilS1 copy1*. Constructs were verified by sequencing. The mc2 region was chosen because gene-conversion events involving this region are well documented, and *copy1* of *pilS1* was chosen because evidence exists for a role of the adjacent *Sma/Cla* repeat in pilin gene recombination (Wainwright *et al.*, 1994, *Mol Microbiol* **13**: 75–78). Introduction of the construct into the gonococcal chromosome was selected for by using a chloramphenicol-resistance gene (*cat*) flanked by directly repeated internal resolution RES sites of the transposon Tn3. After chromosomal insertion, the *cat* cassette was excised by chromosomal integration of the resolvase (*tnpR*) gene of transposon Tn3 into the chromosomal *iga* gene. Finally, by allelic exchange, *tnpR* was replaced with a kanamycin-resistance gene. Gene-conversion events were monitored on the basis of colony-morphology changes which are coupled to P⁺-to-P⁻ transition events. The events were further analysed by PCR, employing primer RB21, which binds specifically to the ribosome-binding site of *pilE* but not to the *pilS* regions, and primer RB22. The *Tetrahymena* intron is spliced from a mRNA precursor (pre-mRNA) both *in vitro* and in *E. coli* (Waring *et al.*, 1985, *Cell* **40**: 371–380). If gene conversion is mediated by a diffusible reverse transcript, the self-splicing event ought to be readily detected because the resulting cDNA would thus lack the intron sequence (Fig. 1). Because the splicing reaction occurs only from plus-strand mRNA, the intron has been inserted in both transcriptional orientations. Although more than 100 P⁺-to-P⁻ switch variants possessing the intron in either orientation were analysed that have undergone gene conversions from *pilS1* to *pilE1*, no clones harbouring spliced introns in *pilE1* were detected. This strongly suggests that

gene conversion in *Neisseria* does not involve a mRNA intermediate of the donor *pilS* gene.

We showed that cell extracts from *N. gonorrhoeae* MS11 were capable of catalysing extensive (>100 bp) reverse transcription on a defined mRNA template. This activity can be explained either by a unique RT present in gonococci or by an intrinsic RT function of DNA polymerase I. The latter explanation is plausible because we showed that purified polymerase I from *E. coli* was similarly active. This interpretation is consistent with a previous report on the RT function of *E. coli* DNA polymerase I based on a different method (Ricchetti and Buc, 1993, *EMBO J* **12**: 387–396). In order to assess whether gonococcal RT activity has a role in gene-conversion processes, we established a genetic system to distinguish between RNA- and DNA-mediated recombination. Using this technique, no evidence for the existence of a RNA intermediate was obtained. This observation suggests that mechanisms other than reverse transcription account for the non-reciprocal transformation-independent recombination events between pilin genes in *N. gonorrhoeae*.

We gratefully acknowledge Scott Gray-Owen for helpful comments on the manuscript.

Roland Barten¹ and Thomas F. Meyer^{1,2*}

¹Max-Planck-Institut für Biologie, Abteilung Infektionsbiologie, Spemannstrasse 34, D-72076 Tübingen, Germany.

²Max-Planck-Institut für Infektionsbiologie, Abteilung Molekulare Biologie, Monbijoustrasse 2, D-10117 Berlin, Germany.

*For correspondence (address no. 2).

E-mail sinfbio@mpib-tuebingen.mpg.de; Tel. (7071) 601 221; Fax (7071) 610 379.

Received 21 February, 1997; accepted 6 March, 1997.

A quest for symbiosis-specific genes urges itself upon *Rhizobium* geneticists

Sir,

Some members of the Rhizobiaceae family are able to oxidize hydrogen (Hup⁺) through regulated expression of hydrogen-uptake (*hup*) genes. This system is particularly useful during symbiosis, when hydrogen is formed by the nitrogenase enzyme complex. Under these conditions, the activity of a hydrogen-oxidizing system is an elegant way of removing hydrogen, which is an inhibitor of the nitrogenase enzyme complex. Interestingly, in Hup⁺ bacteria of the genus *Bradyrhizobium*, the *hup* genes are also expressed in free-living cells, when bacteria are shifted to a micro-aerobic atmosphere containing hydrogen. This might reflect a natural situation, since nitrogenase activity, and thus

hydrogen production, in free-living *Bradyrhizobium* has been reported (Agarwal and Keister, 1983; *Appl Environ Microbiol* **45**: 1592–1601). The transcriptional activator, HoxA, is crucial for expression of *hup* genes in free-living conditions: *hoxA* mutants are Hup⁻ in free-living conditions, but are Hup⁺ in symbiotic conditions, where hydrogenase activity amounts to 50% of the wild-type level (Lambert *et al.*, 1987, *Appl Environ Microbiol* **53**: 422–428; Van Soom *et al.*, 1997, *Mol Microbiol* **23**: 967–977). This implies that HoxA is functioning in bacteroids but can be replaced by a less efficient and so-far unidentified symbiosis-specific factor. On the other hand, Hup⁺ bacteria of the genus *Rhizobium* do not have hydrogenase activity in free-living conditions (Palacios *et al.*, 1990, *Mol Gen Genet* **221**: 363–370). *hup* gene organization in *Rhizobium leguminosarum* biovar *viciae*, is strikingly similar to

that in *Bradyrhizobium*, except that the regulatory genes that border the structural and accessory *hup* genes are missing (*hupUV* to the 'left', *hoxA* and *hupT* to the 'right'). Both the *Bradyrhizobium japonicum* and *R. leguminosarum hupSL* promoter regions contain a well-conserved integration host factor (IHF)-binding site and an RpoN-dependent consensus promoter sequence (Hidalgo *et al.*, 1992, *J Bacteriol* **174**: 4130–4139; Black and Maier, 1995, *Mol Microbiol* **16**: 405–413). In *R. leguminosarum*, the DNA region –172/–87 relative to the start site of transcription is essential for promoter activity, indicating that this area contains the upstream activating sequence (UAS) of a transcriptional activator (Brito *et al.*, 1995, in *Nitrogen Fixation: Fundamentals and Applications*. Tikhonovich *et al.* (eds). Dordrecht: Kluwer Academic Publishers, p. 394). *hup* genes are expressed in nodules induced by *Hup*⁺ *R. leguminosarum* bv. *viciae* but the nature of the transcriptional activator remains unknown. NifA-dependent expression of the *Rhizobium hupSL* genes has been demonstrated in *Klebsiella pneumoniae* hosts, suggesting co-regulation of *nif* and *hup* genes in *R. leguminosarum* (Brito *et al.*, 1995, *ibid.*). As the *R. leguminosarum* bv. *viciae hupSL* genes are not expressed in microaerobically grown cells, the *R. leguminosarum* NifA alone apparently cannot activate *hupSL* expression. It appears that, during evolution, the *hup* structural and accessory genes were highly conserved because of a selective advantage, but the regulatory elements that control the expression of these genes in different bacteria were far less conserved, or even deleted. Activation of *hupSL* expression by the transcriptional regulator HoxA is absolutely required to obtain hydrogenase activity in free-living *B. japonicum* cells, whereas in bacteroids activation of *hupSL* expression can be or has been taken over by a symbiosis-specific factor in *B. japonicum* or *R. leguminosarum*, respectively.

When we compare the nucleotide sequence of the promoter region of the *hupSL* genes in *Bradyrhizobium* and *Rhizobium* using the NMAPUS program (PC-GENE, IntelliGenetics), it is striking that a conserved CT-rich region, with 10 out of 12 identical nucleotides, is present in the upstream regulatory regions that are essential for *hupSL* promoter activity (Fig. 1). Whether this conserved site is the UAS for the symbiosis-specific factor remains to be demonstrated.

```

Bj UAS      -131 TTTCCGCTTCTC      -120
                | | | | | | | |
Rl UAS      -112 TATCCGGTTCTC      -101

```

Fig. 1. Nucleotide sequence comparison (NMAPUS; PC-GENE, IntelliGenetics) of part of the upstream regulatory region of the *B. japonicum* (Bj UAS) and *R. leguminosarum* (Rl UAS) *hupSL* genes. Numbers are relative to the start site of *hupSL* transcription. The –149/–98 region in *B. japonicum* and the –172/–87 region in *R. leguminosarum* are required for *hupSL* transcription (for references, see the text).

The differential regulation of *hup* genes in free-living and symbiotic rhizobia could be explained according to the 'regulatory noise' hypothesis (de Lorenzo and Pérez-Martin, 1996, *Mol Microbiol* **19**: 1177–1184), which states that structural and regulatory genes evolve relatively independently and that new transcriptional control systems can evolve because of the non-specificity of pre-existing promoters. It is well known that the differentiation process leading to an intracellularly nitrogen-fixing bacteroid is accompanied by a drastic change in bacterial metabolism and morphology. This implies that different, bacteroid-specific regulatory factors act on bacteroid-specific genes. In the regulatory noise hypothesis, these alternative regulators could interact in an illegitimate way with pre-existing promoter sequences of genes that are normally regulated by other transcriptional factors. When this illegitimate expression has a selective advantage, the newly recruited, sub-optimal transcriptional control unit will be fine-tuned in order to respond predominantly to the new signal.

Another example of differential regulation in free-living versus symbiotic cells can be found in the regulation of *dctA*, encoding a dicarboxylate permease in *Rhizobium meliloti*. The transcriptional activator DctD is required for RpoN-dependent expression of the structural genes in free-living cells but can be replaced in bacteroids by a symbiosis-specific regulator, which is neither NifA or NtrC (Ronson *et al.*, 1987, *J Bacteriol* **169**: 2424–2431; Yarosh *et al.*, 1989, *Mol Microbiol* **3**: 813–823; Allaway *et al.*, 1995, *FEMS Microbiol Lett* **128**: 241–245). Differential regulation, albeit more subtle, also exists in the case of *nifHDK* expression in *R. meliloti*: the same regulator (NifA) binds at different *cis*-elements in the free-living versus bacteroid state (Better *et al.*, 1985, *EMBO J* **4**: 2419–2424; Ditta *et al.*, 1988, in *Molecular Genetics of Plant–Microbe Interactions*. Palacios and Verma (eds). St. Paul, Minnesota: APS Press, pp. 109–110; Wang *et al.*, 1991, *Mol Microbiol* **5**: 157–162). The conformation of the UAS and/or of the regulatory protein might be influenced by the symbiotic state, leading to differential affinities of the regulatory protein for certain UASs.

The identification of hitherto unknown factors responsible for symbiosis-specific expression will be of great interest, as it is becoming increasingly apparent that studies on symbiotic regulatory processes *ex planta* reflect a virtual reality which can be different from real life in the nodule.

Carolien Van Soom and Jos Vanderleyden*
*F. A. Janssens Laboratory of Genetics, K.U. Leuven,
 W. de Croylaan 42, B-3001 Heverlee, Belgium.*

*For correspondence.

*E-mail janssens@agr.kuleuven.ac.be; Tel. (16) 322403;
 Fax (16) 322990.*

Received and accepted 27 February, 1997.