Construction and Utilisation of a Bidirectional Reporter Vector in the Analysis of Two *nod*-Boxes in *Rhizobium loti*

A thesis presented in partial fulfilment of the requirements for the degree of
Masters of Science
in Molecular Genetics at Massey University

Simon Keith Parry
1993

ABSTRACT

The *nod*-box is a 47bp *cis*-acting regulatory region which has been conserved amongst every species of *Rhizobium* studied to date. In species such as *R. meliloti* and *R. leguminosarum*, the *nod*-box has been shown to promote constitutive activity towards the regulatory *nod*D gene, and flavonoid-inducible expression towards the divergently-transcribed *nod*ABCIJ operon. This bidirectional regulation of the so-called common *nod* genes was not observed in *R. loti*. A previous analysis of this species had shown that its *nod*-box promoted inducible activity towards the truncated '*nod*D' gene, as well as the *nod*ACIJ operon. It was the unusual arrangement of these *R. loti nod* genes that had initially aroused interest in this bacteria.

To further investigate the role of the *nod*-box in the regulation of the *R. loti* common *nod* genes, a bidirectional reporter vector (pSPV4) was constructed. This novel vector allowed the promoter activity of a cloned *nod*-box-containing fragment to be concurrently measured in either direction using the same culture of cells. To achieve this construct, the *gus*A gene from pRAJ260 was blunt-end ligated into pUC21. An in-frame ribosome binding site (rbs) was cloned upstream of the *gus*A coding sequence to facilitate transcriptional fusions. The rbs and *gus*A gene were later excised as a functional unit and blunt-end ligated into pMP220 alongside the ß-galactosidase reporter gene but in the opposite orientation. Hence, both reporter genes could be divergently transcribed from a common regulatory region cloned into the multiple cloning site that separated the genes.

The fragments of DNA that were eventually cloned into the bidirectional vector were generated through the polymerase chain reaction. Each DNA insert contained the *nod*-box bracketed by differing lengths of flanking region. Once these PCR-generated fragments had been sequenced in pUC118 and subcloned into pSPV4, the resulting constructs were transformed into *R. loti* cells by electroporation. As the electroporation of these cells had not previously been reported, the conditions for this procedure were established and optimised.

The results obtained from the bidirectional reporter assays disagreed with those observed in the earlier assays by Teo (1990). Neither the *nod*ACIJ *nod*-box of NZP2037 nor the *nod*B *nod*-box of NZP2213, showed bidirectional inducible expression. In fact, both *nod*-boxes showed constitutive expression in the '*nod*D' direction and inducible expression in the opposite direction. This indicates that the control of the *nod* genes in *R. loti* is fundamentally the same as that seen in other fast-growing *Rhizobium* species. Three regulatory elements affecting the levels of *nod* gene expression have tentatively been identified outside the *nod*-box sequence, though the results indicating their presence may simply be due to spacing differences between the *nod*-box and the reporter gene.

ACKNOWLEDGEMENTS

Primarily, I should like to thank my chief supervisor Associate Professor Eric Terzaghi for his advice, encouragement and patience at all stages of the research. Hopefully, the inexplicable experimental results that I regularly presented him with did not contribute too much to his early 'retirement'. Similarly, sincere thanks must go also to my co-supervisor Shashi Sharma (CRI AgResearch Grasslands) who conceived the notion of a bidirectional reporter vector and provided much assistance in its development.

Thanks must also go to the Department of Microbiology and Genetics for providing the facilities and funding necessary for the project to be completed. I am also indebted to Professor Barry Scott for the bacterial strains he kindly supplied, and for his willingness to provide advice throughout the project. Acknowledgement is also given to Drs Brian Mansfield, Rosie Bradshaw and Hale Nicholson for their discussions and strategic advice at various stages of the work. Thanks also goes to Trish McLenachan and Carolyn Young for technical guidance.

Paul Hirst and David Harte helped to keep me going during the more difficult stages of my work and their support was much appreciated.

My parents have also been a great source of support, both emotionally and financially. Their help has been has been gratefully acknowledged.

Finally, thanks must go to my girlfriend Rhonda Stanway for her love, friendship, proof-reading ability and cookery skills she demonstrated throughout the course of this project. Rhonda's patience and tolerance were also greatly appreciated.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 - INTRODUCTION

1.1 SYMBIOTIC RELATIONSHIP BETWEEN *Rhizobium* AND LEGUMES

In many communities the most stable and persistent associations between species are those based on obligative mutualism. Since the interaction is mandatory for their survival, the existence of both species in a given habitat is essential. Because of this requirement, a variety of intricate structural or behavioural adaptations have evolved whereby the partnership is maintained throughout the generations. In other cases the relationship is perpetuated simply through the high probability of mutual encounters resulting from a substantial population density or vast offspring numbers contributed by one or both members.

There is a well-characterised mutualistic relationship between *Rhizobium* and leguminous dicotyledons. In this case the *Rhizobium* stimulates the legumes to develop root nodules, which the bacteria infect and inhabit. Ultimately the two symbionts cooperate via metabolic exchanges, enhancing their mutual survival. More specifically the bacteria reduce atmospheric nitrogen into ammonia which is then exported to the plant for assimilation into protein and other nitrogenous compounds. The plant on the other hand supplies photosynthates which provide the energy necessary for bacterial multiplication and nitrogen fixation. Studies of mutants have shown that the major carbon-based energy sources exported by the plant to the bacteroids are dicarboxylic acids, such as succinate, malate and fumarate (Ronson *et al.*, 1981; Engelke *et al.*, 1987; Yarosh *et al.*, 1989). Salminen and Streeter (1992), using C^{14} labelling, were able to demonstrate that a significant proportion of the carbon imported into the *Rhizobium leguminosarum* bacteroids was diverted into the production of glutamate via the tricarboxylic acid (TCA) cycle, and the fixed nitrogen exported to the plant in this form.

Although the aforementioned mutual benefits are considerable, they are not strictly obligatory. In fact, both parties to the symbiosis may exist independently of the other. The *Rhizobium* can be found as a free-living soil organism though its existence is more sensitive to external environmental factors than its nodule-dwelling counterparts. Similarly, nodulation is not always imperative for the survival of legumes in nitrogen-rich soil, but it does appear to confer selective advantages. This is borne out by the observation (reviewed by Beringer *et al.*, 1979), that leguminous seeds contain a significantly higher protein content than the seeds of non-legumes (25% compared to 10% in wheat). Conceivably the extra requirement of seeds for nitrogen may have been an important factor in the selection for nitrogen-fixing nodules in leguminous plants.

1.2 *Rhizobium*-INDUCED NODULATION OF LEGUMINOUS PLANTS

An important part of the symbiotic relationship between *Rhizobium* and the legumes is the initiation of the formation of the nodule. The process that causes the plant to undergo the transition from normal root growth to the development of nodules is fascinating in that it represents a situation where the *Rhizobium* exerts a regulatory influence on the plant genome. This results in the differentiated meristematic tissue becoming developmentally active again, and redifferentiating into nodule tissue. If this important regulatory process could be understood, it would be of great use in gene manipulation studies, both in a pure and applied context.

1.2.1 Brief Summary of Nodule Formation

The development of nitrogen-fixing nodules may be attributed to a series of complex interactions between the legume and compatible rhizobia (Newcomb, 1981). In general, the sequence of events can be likened to a two-way conversation between the plant and the bacteria.

1.2.1.1 Chemotaxis of Rhizobia to Legume Roots

The movement of the rhizobia to the legume roots is an important first step in the formation of a nitrogen-fixing nodule. Interestingly, a number of other organisms, such as *Azospirillum*, exhibit motility towards roots due to specific attractants diffused therefrom (Bashan and Lavanony, 1987). It is very likely that the rhizobia follow a similar form of dispersal as they have been shown to be motile and chemotactic towards nutrients (Bowra and Dilworth, 1981; Bergman *et al.*, 1988) and root exudates (Gitte *et al.*, 1978; Caetano-Anolles *et al.*, 1988).

The attraction exhibited by the *Rhizobium* in response to nutrients enables the bacteria to reach new sources of energy substrate. Indeed it has been reported that wild-type *R. leguminosarum* bv. *trifolii* grew considerably better than non-motile mutants of the same variety when inoculated into sterile soil (Barnet, 1991). Clearly the wild-type organisms gained a selective advantage over their non-motile and non-chemotactic competitors due to their ability to migrate to improved nutritional environments.

Leguminous roots release secondary plant metabolites, such as rhizopines, into the immediate surroundings of the rhizosphere, where the wild-type rhizobia are able to utilise these products as a carbon and energy source (Murphy *et al.*, 1987). One recently investigated metabolite, trigonelline, has also been implicated as an energy source for the rhizobia in both the rhizosphere and the nodule itself (Boivin *et al.*, 1990). In this case

the *trc* genes found on the *Rhizobium* pSym plasmid are induced by the trigonelline to produce proteins that can catabolise the plant product. The evolution of inducible genes able to catabolise these specific plant products has conferred enormous competitive advantages over other bacteria.

As mentioned earlier, nutrients are not the only compounds to which rhizobia are attracted. Root exudates also elicit a positive chemotactic response. Caetano-Anolles *et al.* (1988) and Kape *et al.* (1991) have identified specific flavonoid compounds contained within the root exudate as being the chemical attractants. Presumably the combined chemotactic pressure exerted by the plant-derived nutrients and flavonoid compounds draws the rhizobia towards the leguminous roots.

1.2.1.2 Induced Rhizobial Response to Leguminous Flavonoid Compounds

The flavonoid compounds also have a regulatory role to play in the formation of the nodule. They are indirectly responsible for regulating the expression of the *Rhizobium* nodulation (*nod*) genes. The actual process by which the *nod* genes are controlled appears to be a departure from the phage λ paradigm of gene regulation and constitutes the primary subject of investigation for this thesis. The *nod* genes code for a variety of enzymes which are involved in the synthesis of Nod factors (Redmond *et al.*, 1986; Peters and Long, 1988). These morphogens play an important role in inducing the formation of nodule primordia in the root cortex.

The chemical structure of Nod factors synthesised by several *Rhizobium* strains has recently been examined and analysed. The basic conformation of these factors appears to be a ß-1,4-linked oligomer of N-acetylglucosamine bearing an N-acyl substitution on the non-reducing end (Lerouge *et al.*, 1990; Truchet *et al.*, 1991; Spaink *et al.*, 1991). Slight differences in the length of the oligomeric backbone and the fatty acyl substitutions result in a large family of possible Nod factors. Further variation in the Nod factors can be observed at the reducing end of the compound. This terminus may be modified by the addition of either a sulphate, fucose or methylfucose group depending on the species. The minor differences between the families of Nod factors synthesised by each rhizobial strain are thought to be at least partially responsible for host range specificity. Indeed Roche *et al.* (1991), and Spaink *et al.* (1991), have shown that these modifications are critical for the host-specific induction of nodule primordia. The idea that simple chemical structures can induce profound developmental changes is not new. Albersheim and Darvill (1985), working extensively with oligosaccharides demonstrated that these simple compounds can act as regulatory molecules at very low concentrations.

The synthesised Nod factors are released from the *Rhizobium* and constitute a molecular response to the plants initial flavonoid signal. The diffusible properties of the Nod factors were examined by Ames and Bergman (1981), and Mellor *et al.* (1987). They found that non-motile mutants retained the ability to induce nodulation. It therefore follows that the movement of the rhizobia towards the legume is not essential for nodulation to occur. The message from the bacteria to the plant is instead delivered by the diffusible Nod factor.

The means by which the Nod factor signals the plant to form a nodule is not well understood. Recently, however, Ehrhardt *et al.* (1992), measured transmembrane potential change in single infectible root hair cells in response to *Rhizobium meliloti* extracellular Nod factor NodRm-IV(S). It was found that a rapid depolarisation of transmembrane potential occurred when the Nod factor was present. A second messenger, such as auxin or cytokinin, may be responsible for transducing the signal for nodule formation to the root interior (Hirsch *et al.*, 1989). Perhaps depolarisation of the membrane and changes in plant hormone levels are somehow involved in a Nod-factor-induced signal cascade. A considerable amount of further experimentation will be required before this facet of nodulation can be elucidated.

1.2.1.3 Formation of the Nodule

The rapidly expanding root hair cells are susceptible to *Rhizobium* Nod factor-induced deformation (Bhuvaneswari *et al.*, 1981). Although the curling can be induced by bacteria-free lysates, it is most severe when the rhizobia is bound. Van Batenburg *et al.* (1986), suggests this process can be attributed to the *Rhizobium* provoking a local stimulation in the rate of plant wall expansion.

The *Rhizobium* and the root hair are able to maintain a tight attachment to one another through the formation of cellulose fibrils. These compounds assist in embedding the *Rhizobium* into the mucilaginous material protecting the surface of the hair cells. Furthermore, the binding may be strengthened by some calcium-dependent proteins (adhesins) produced by rhizobia (Smit *et al.*, 1989). Lectins produced by the host bind to both the plant and compatible rhizobia. Dazzo and Gardiol (1984), and Diaz *et al.* (1989), proposed that this lectin-mediated attachment may be an important step early in the recognition of rhizobial strains that can form a nodule. This hypothesis has been reinforced by recent experiments by Philip-Hollingsworth *et al.* (1989), working with *R. leguminosarum*. When the microsymbiont *R. leguminosarum* 300 received the *nod* genes specifying the host range of the clover microsymbiont *R. leguminosarum* ANU843, its *in situ* binding of clover lectin (trifoliin A) increased greatly to a level comparable to that of the clover symbiont ANU483. Furthermore, when Diaz *et al.* (1989), transformed

white clover with pea lectin genes, they observed that the clover's roots were susceptible to nodulation by the pea microsymbiont 248.

The attachment of the bacteria to the plant root elicits a sophisticated developmental response in the legume beyond that of root hair curling. This deformation coincides with the initiation of cell division in the sub-epidermal layer of the root cortex (Newcomb, 1981; Dudley *et al.*, 1987; Guinel and LaRue, 1991). The activation of the root cortical cells leads in turn to the formation of the nodule meristem, whose growth pattern is vastly dissimilar from the neighbouring lateral root cells at the same temporal stage (Dudley *et al.*, 1987). The infected plant cells are then stimulated to produce cell wall sheathes termed 'infection threads' (Callaham and Torrey, 1981). The infection thread, which can be descibed as an inwardly growing tunnel, invades a cell in the outermost layer of the root cortex before infecting subsequent cell layers in a pre-determined manner. These plant-derived threads provide access for the rhizobia into the nodule primordia. When the bacteria have reached the end of the sheath they get released into the plant cells by endocytosis (Dart, 1977; Goodchild, 1978).

At this stage of nodule development the rhizobia become enveloped in host-derived peribacteroid membranes (Robertson *et al.*, 1978; Bauer, 1981), which are antigenically similar to the plant cell plasma membrane (Perotto *et al.*, 1991). The intracellular rhizobia then differentiate into pleiomorphic forms called bacteroids which are able to fix nitrogen and exchange metabolites with the differentiated nodule tissue (Robertson *et al.*, 1978; Verma and Long, 1983).

The formation of nodules is accompanied by the nodule-specific expression of the plant nodulin genes. Early nodulin genes are expressed when the plant becomes infected and the nodule begins to form, whereas the more numerous late nodulin genes are first expressed around the onset of nitrogen fixation (reviewed by Franssen *et al.*, 1989).

1.3 THE NODULATION (*nod*) GENES OF *Rhizobium*

The rhizobial genes responsible for inducing the plant to undergo the transition from normal root growth to the initiation of nodules are the nodulation (*nod*) genes. Two different classes of *nod* genes have been identified in *Rhizobium*, both of which are required for the nodulation of the specific host legume.

The first group of nodulation genes, called the common *nod* genes, are so named because they appear to be functionally interchangeable between rhizobia. Mutations in one *Rhizobium* species can be complemented by cloned wild-type genes from another

species. Furthermore, these cloned genes are insufficient to extend the host range of any recipient species. The best characterised common nodulation genes found in *Rhizobium* are the *nod*ABC and *nod*D genes. Mutations in the *nod*DABC genes of *R. leguminosarum* bv. *viciae* (Downie *et al.*, 1985) and bv. *trifolii* (Djordjevic *et al.*, 1987) resulted in legume roots exhibiting neither hair curling (Hac$^-$) nor nodule formation (Nod$^-$). These genes are also necessary for the induction of cortical cell division in the plant root (Dudley *et al.*, 1987).

Based on the observations that the *nod*ABC genes are required in these pivotal steps of nodule formation, Long (1989) suggested a common antecedent for these events. Support for this proposal came from Marvel *et al.* (1987). They showed that the *nod*ABC gene products were even required in the nodulation of *Parasponia*, the only non-legume genus whose members form *Rhizobium*-induced nodules.

Biochemical and immunological studies have shown that NodA and NodB are localised in the cytosol (reviewed by Denarie *et al.*, 1992). It is thought that these cytosolic proteins may be involved in the synthesis of the amino sugar backbone of the Nod factor. Alternatively, the NodAB proteins may have a role in the N-acylation of this signal molecule.

The other common *nod* gene involved in the production of the Nod factor is *nod*C. Sequence comparisons have revealed that the *nod*C gene product shows significant homology with chitin synthases from *Saccharomyces cerevisiae* (Bulawa and Wasco, 1991). This suggests that NodC is the major protein responsible for synthesising the chitin oligomeric backbone of the Nod factor. Immunological studies have demonstrated that at least part of the NodC protein is located in both of the two bacterial membranes (John *et al.*, 1988; Hubac *et al.*, 1992). This indicates that it could also be involved in the transport of the Nod factor out of the cell.

The common *nod*ABC genes are all under the control of the NodD protein. This positive transcriptional activator has been localised exclusively to the cytoplasmic membrane in *R. leguminosarum* bv. *viciae* (Schlaman *et al.*, 1989). Schlaman *et al.* (1989), proposed that the amphipathic NodD may be anchored in the cytoplasmic membrane with a substantial hydrophilic domain extending into the cytoplasm. The structure and regulatory properties of NodD will be discussed in far greater detail in a following section. It will also become evident that while *nod*D has been classified as a common *nod* gene on the basis of cross-complementation studies, it is also involved in host recognition. At this stage it should merely be noted that the *nod*D gene product mediates the first step in the plant-microbe dialogue by means of interactions with a plant-specific flavonoid (as discussed earlier).

The other common nodulation genes are *nod*I and *nod*J, about which little is known. It has been suggested that their coded proteins form a transport complex in the membrane of the *Rhizobium* (Evans and Downie, 1986).

The second group of nodulation genes consists of the host-specific nodulation (*hsn*) genes. This class is involved in the decoration of the Nod factor and is therefore responsible for the recognition of potentially nodule-bearing host plants. Evidence for this is shown by mutations in the *hsn* genes, which result in expanded host ranges (Kondorosi and Kondorosi, 1986). The entire *hsn* group of genes are unable to complement *nod* mutants where the mutation is located within one of the *hsn* genes from heterologous *Rhizobium* species, suggesting high species specificity. There are also some genes, within this class (including *nod*EF, *nod*H and *nod*LMN), which affect the position and tightness of root haircurling, and the regulation of cell division (reviewed by Long, 1989).

1.4 PHAGE LAMBDA (λ) REGULATORY PARADIGM

Before examining the regulation of the *nod* genes in detail, it would be of great benefit to outline the general features of the λ regulation system. This system illustrates many aspects of gene control that are thought to be essential in the promoter/operator function of almost all the genes studied to date. It is, therefore, a valuable model with which the regulation of the *nod* genes may be compared and contrasted.

1.4.1 Overview of Phage Induction

The bacteriophage λ is a very well studied virus whose biochemistry and genetics have been exhaustively examined and understood. Included in this analysis has been the dissection of the regulatory control features exhibited by the phage. Now, through comparisons with regulatory systems found in other organisms, these features have emerged as a useful model of promoter/operator systems.

Bacteriophage λ possesses two modes of replicating its DNA. In physiologically favourable conditions the λ DNA is incorporated into the DNA of its host where it is passively replicated along with the other elements of the bacterial chromosome. However, when an infected bacterium is exposed to unfavourable conditions, such as ultraviolet (UV) light, a different train of events is set in motion. Once the phage has sensed that the external disturbance is likely to compromise the bacterium, it abandons its previously useful strategy of lysogenic existence and enters the lytic pathway. This mechanism of switching from the lysogenic cycle to the lytic cycle involves the promoter/operator control system, long considered the paradigm of control function. It is the objective of this discussion to describe this paradigm with a view to comparing the superficially similar mode of bidirectional control exhibited by the *nod*-box in *Rhizobium*, at a later stage.

1.4.2 Components of the Genetic Switch

The regulatory proteins responsible for controlling the switch are the λ repressor and the Cro protein which are encoded by the phage genes *c*I and *cro* respectively. These two genes, separated by 80 bp, are transcribed divergently from promoters located in this intergenic region. The *c*I promoter, called P_{RM}, directs the bacterially-encoded RNA polymerase to the left, whereas the *cro* promoter, P_R, orients the polymerase in the opposite direction. Significantly, the *c*I and *cro* promoters are juxtaposed yet do not overlap (Figure 1.1).

Figure 1.1. Genetic Organisation of the Phage λ Control Circuitry

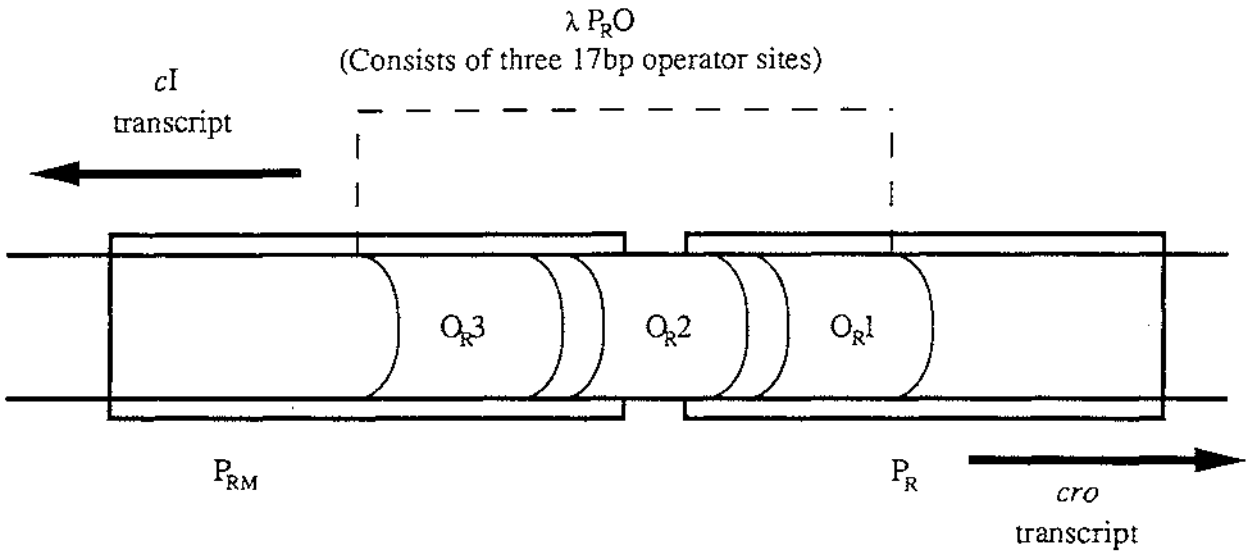


Figure 1.1

The *cis*-acting genetic components of the phage λ regulatory switch. O_R1, O_R2 and O_R3 represent the tripartite operator to which the regulatory proteins bind. Overlapping these sites to varying degrees are the two divergent promoters, P_R and P_{RM}, which provide the RNA polymerase binding sites allowing transcription of these proteins. (After Ptashne).

The 80bp intergenic segment contains another regulatory site in addition to the divergent promoters. This is the right operator of λ, which is comprised of three adjacent protein binding sites (O_R1, O_R2 and O_R3). Each of these 17bp operator sites functionally overlaps either one or the other promoter, or in the case of O_R2, both of the promoters. Sequence analysis has shown that the three operator sites are all imperfect inverted repeats, which are similar to one another. The inverted repeat symmetry is thought to be structurally important in the binding of dimeric proteins.

Although the three protein-binding sites display partial sequence symmetry about an axis through the central base pair, they are not identical. The regulatory proteins can distinguish between them. For example, considering any two operator sites, one might have a higher affinity than the other for Cro. Hence at a specific concentration, it is more likely that a Cro molecule would be found at the site with a higher binding constant, than at the site with a lower binding constant. A hierarchy of affinity can be elucidated for each of the regulatory proteins (Cro and the λ repressor) and for each of their three binding sites (O_R1, O_R2 and O_R3).

The λ repressor protein can exist in a monomeric or dimeric form. Formation of the functional dimeric repressor structure is facilitated by interactions between carboxyl domains of individual monomers. The energetic stability afforded by the dimeric configuration results in the predominance of this favourable species in the lysogen. The other functional unit of the repressor protein is the N-terminal domain which is linked to the C-terminal domain by a 40 amino acid residue connector. While the C-terminal domain of the repressor is responsible for the formation of dimers, the N-terminal domain has a DNA-binding role. Unsurprisingly, this domain contains recognition sites that bind specifically to the operator sites on the integrated phage λ chromosome.

Previous *in vitro* investigations have shown that both types of repressor monomeric domain can exercise their function independently of the other. Firstly, the dissociated N-terminal domains are still capable of binding to DNA operator sites, yet do so less readily than the intact repressor molecule. This difference in affinity is due to the cooperative effect conferred by the dimeric form for inverted repeats. The repressor dimer is thought to induce a conformational change which results in this form having a higher binding constant than the monomeric form (see Figure 1.2a). The C-terminal domains behave in a similar fashion albeit without cooperativity. In the absence of N-termini, these domains are able to form oligomeric chains *in vitro*. This demonstrates that the inter-monomeric binding capacity lies solely in this C-terminal region but does not comment on their *in vivo* activity.

The second regulatory protein, Cro, is considerably smaller than the λ repressor (66 aa compared to 236 aa). In fact the Cro monomer folds into a single globular conformation whose size is similar to the amino domain of the repressor. As with their regulatory counterparts, the Cro monomers have an equilibrium constant that favours the formation of dimers. Predictably, these Cro dimers also exhibit a much greater affinity for DNA operator sites than the monomeric form.

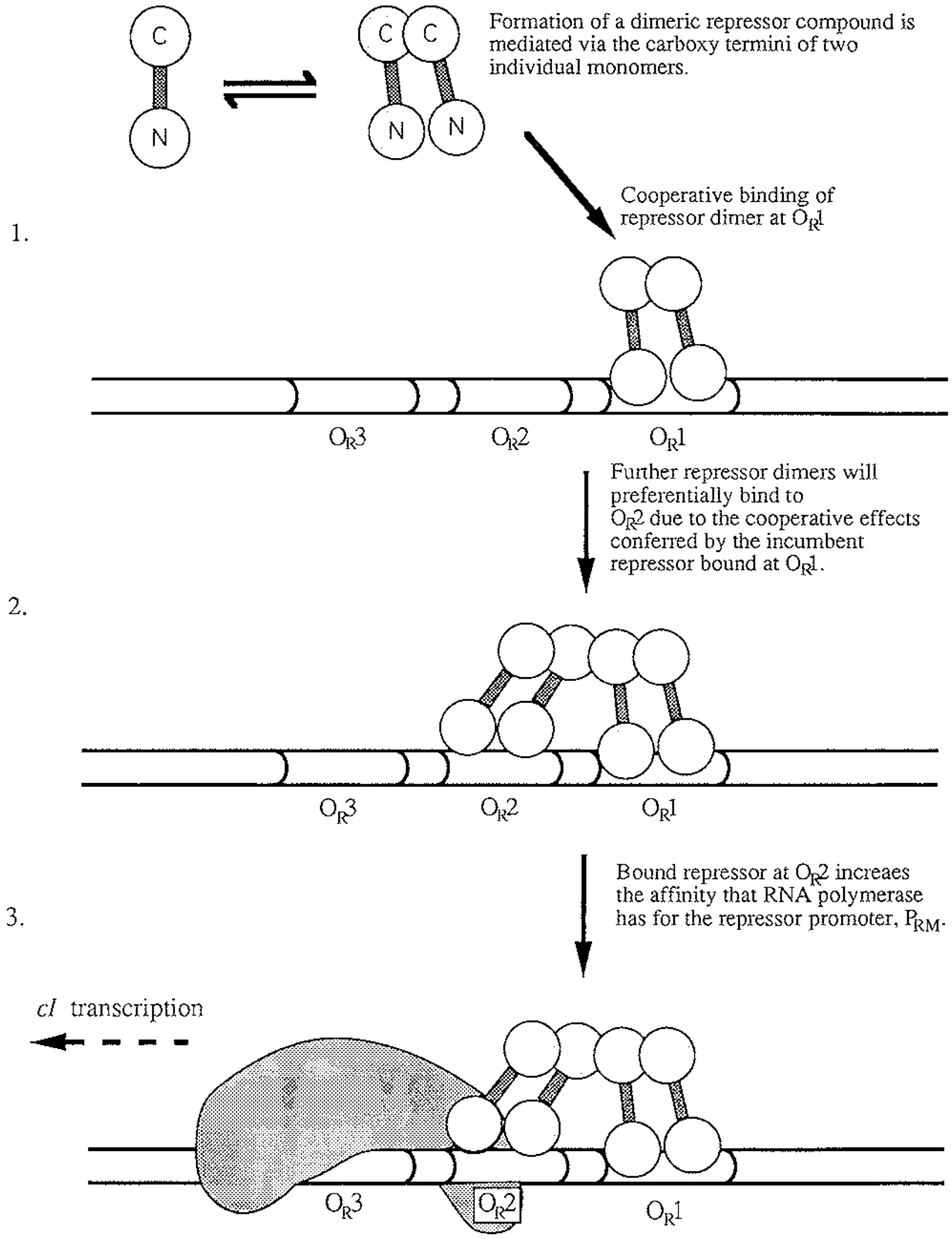
1.4.3 The Lysogenic State

In favourable conditions, it is important for the phage to maintain the lysogenic state. Being the only phage-encoded protein synthesised during this passive phase, the repressor molecule is the primary determinant of its continuation. To effectively maintain lysogeny the concentration of the repressor dimers must be regulated to ensure levels remain within a window of physiological functionality. If the concentration of the repressor dimers drop to an extremely low level, such that P_R is unblocked, premature induction of the lytic phase would certainly follow. On the other hand, if the repressor becomes greatly overexpressed, the phage would be insensitive to external stimuli that would usually induce cell lysis.

Appropriate levels of repressor protein are maintained through an elegant regulatory system. When faced with the triplication of operator binding sites at the λO_R, the repressor dimer binds to the site to which its affinity is the greatest. In this case it is always site 1. The preferential occupation of this site is an important feature in the maintenance of the lysogenic state. A cooperative interaction is thought to be mediated via protein-protein interactions between the bound repressor molecule and a second incoming dimer. This interaction results in the repressor dimer at O_R1 greatly increasing the affinity with which a second dimer can bind to site 2. The connector region of the first repressor is thought to orient the C-terminal regions of the dimer in such a way that they contact the C-terminal domains of the second dimer (see Figure 1.2(b)). This cooperative effect cannot be extended to a third repressor dimer binding to O_R3 because of steric limitations. The lack of a cooperative effect, taken with the low affinity of the repressor dimer for O_R3, means that site 3 is usually unoccupied.

A second effect of dimers binding to O_R1 and O_R2 is one of positive control. Initial contact between the N-terminal domain of the repressor molecule bound to O_R2 and the bacterial RNA polymerase, assists the latter to bind to P_{RM}. This cooperative effect allows the RNA polymerase to preferentially transcribe additional *cI* mRNA molecules. Hence, it would appear that the occupation of these two operator sites by repressor proteins stimulates transcription of their own *cI* gene (see Figure 1.2(c)).

Figure 1.2. Cooperativity Exhibited By the λ Repressor



The binding of repressor dimers to O_R1, and latterly to O_R2, displays negative regulatory control in addition to the positive control just outlined. Indeed, the regulatory suppression of the *cro* transcript by the repressor dimers is an integral feature of the λ switch. The negative control is mediated through the superposition of O_R1 and O_R2 with the Cro promoter, P_R. The binding of these sites by repressor molecules prevents the bacterial RNA polymerase from transcribing the *cro* gene. Therefore, by blocking the P_R promoter, the repressor dimers are effectively abolishing expression of the lytic genes, whilst concomitantly promoting transcription of their own lysogenic *c*I gene. The position of the λ switch is therefore solely dependent on the maintenance of an optimal repressor concentration.

Sometimes the rate of cell division varies, resulting in fluctuating concentrations of repressor molecule. For the integrated phage to remain sensitive to inducing agents, this repressor concentration must be kept below a specific level. In circumstances where the concentration of repressor exceeds this threshold value, the dimers will occupy the O_R3 site which has a higher binding constant than the other sites. This binding physically blocks the RNA polymerase from proceeding to transcribe further repressor molecules. When the cell divides, there is a resulting decrease in the repressor concentration, binding of the O_R3 site is attenuated, and transcription of further repressor molecules will actively proceed. As mentioned earlier, this transcription is enhanced by the cooperative interaction between the two bound dimers at O_R1 and O_R2 with the adjacently bound RNA polymerase.

1.4.4 Switching from the Lysogenic State into the Lytic Cycle

The lysogenic state of the phage could be maintained indefinitely in the absence of external agents. However, irradiating the lysogen with UV light induces a dramatic change in life cycle. In a process yet to be resolved, the RecA protease activity somehow becomes activated and cleaves the connecting region of the repressor molecules. The separation of the amino domain from the carboxyl domain effectively inactivates the repressor because the separated amino domain cannot dimerise. As the monomers have a lower affinity for the DNA than the dimeric form. It is unlikely that any monomeric amino domains will bind to the operator sites. It therefore follows that when the repressor molecules vacate the operator sites there will be too few uncleaved repressor dimers available to replace them.

The resultant absence of repressor molecules bound to the tripartite operator, leads to two important changes. Firstly, there is a decrease in the amount of repressor protein synthesised. The reason for this decline lies in the self-regulatory properties mentioned earlier. As the concentration of the repressor dimers falls, the cooperative effect between the repressor dimer at O_R2 and the RNA polymerase is lost. The expression of the *cI* gene is therefore no longer stimulated, and a negative cascade occurs.

Further disruption of the lysogenic cycle is caused by the initiation of transcription of the previously unexpressed phage genes. Of these, a second phage regulatory protein, Cro, is synthesised, and binds to the vacant operator sites. In a situation opposite to that exhibited by repressor dimers, the Cro dimers have the highest affinity for O_R3. Therefore, once bound, the Cro protein abolishes *cI* synthesis by blocking the polymerase binding site of the repressor. At this point the phage has switched from the lysogenic cycle to the lytic train of events.

1.4.5 Important Features of Phage λ Regulation

The previous sections outlined the series of events which cause phage λ to switch from the lysogenic state to the lytic cycle. The three main points regarding this λ regulatory paradigm are:

1) Operator sites contain inverted repeats which allows dimers to bind, i.e. one monomer subunit per half a binding site.

2) Different affinities for the operator sites by the repressor and the Cro protein results in regulatory proteins binding in a specific order.

3) Cooperativity at several levels enhances binding, and therefore λ regulation.