FEBS 16753

FEBS Letters 382 (1996) 1–5

Hypothesis

Regulation of gene expression at a distance: the hypothetical role of regulatory protein-mediated topological changes of DNA

Ferenc Marincs*, Derek W.R. White

Plant Molecular Genetics Laboratory, AgResearch, Grasslands Research Centre, Tennent Drive, Private Bag 11008, Palmerston North, New Zealand

Received 12 December 1995; revised version received 31 January 1996

Abstract A theoretical model is presented that a regulatory protein may activate the transcription of a promoter by interacting with a single remote operator. In response to an inducer molecule the regulatory protein bound to the operator undergoes a conformational change, and might mediate a B to Z-DNA conversion of the operator. This transition would remove both helical turns and supercoils from the intervening region between the operator and the promoter, resulting in the correct spatial arrangement of the −10 and −35 hexamers of the promoter, which therefore can be efficiently transcribed.

Key words: Gene expression; Remote operator; Z-DNA; Supercoiling

1. Introduction

Regulation of gene expression in bacteria is mediated by DNA-binding proteins, whose interaction with operator sequences results in repression or activation of the regulated promoters. The position of the operator relative to the regulated promoter is defined as proximal, if the operator is situated between the −65 and +20 positions of the promoter, or remote, if the operator is lying outside of the −65 to +20 region [1]. The manner in which the binding of a regulatory protein to its operator site affects transcription from the regulated promoter is an important theme in molecular biology. In operons where the position of the operator is proximal or slightly remote, gene expression is repressed or activated by direct contact between the regulatory protein and the RNA polymerase [2]. However, the mode of regulation is less obvious in situations in which the operator has an extreme remote position, and therefore no direct contact is possible between the regulatory protein and the RNA polymerase for regulating transcription.

To date three basic models have been presented, which describe the molecular mechanism of gene regulation at a distance. In the first model, a regulatory protein binds cooperatively to differently positioned operators causing loop formation between the operators [3]. This type of regulation is known in many bacterial systems; examples include the λ repressor, and the *are*, *deo*, *gal* and *lac* operons of *Escherichia coli* [4]. These systems, however, require that both remote and proximal operators are involved. In the second model, a regulatory protein bound to a remote operator interacts with the RNA polymerase bound to the distant promoter, which causes loop formation in the intervening DNA, and results in the activation of the promoter [4,5]. The NtrC and XylR proteins appear to activate transcription by this mechanism [6,7]. In a third model it is visualised that a regulatory protein bound to a remote operator site mediates expression of the distant promoter by modulating the conformation of the DNA between the operator and the promoter [8]. In contrast to the first two models, there are no examples for the third model to date, and in this paper we have been exploring the possibility of this type of regulation in a bacterial operon.

2. Possible conformational alterations of the DNA

In the third model briefly outlined above, it is presumed that the interaction between a regulatory protein and a remote operator induces a conformational change in the DNA double helix toward the distant promoter, and that such a conformational change turns on or off the promoter. The most important question, according to this model, is to determine the type of conformational change that could be induced along the DNA helix by the bound regulatory protein, which in turn leads to regulation of the promoter at a distance. Since this model excludes both the presence of an operator at the promoter site and direct contact between the regulatory protein and the RNA polymerase (important facets of the first two models), no looping of the intervening region is conceivable.

Bending of DNA by the bound regulatory protein, although previously described for a number of bacterial regulators, would not seem to be sufficient to regulate the expression of a distant promoter, because this mechanism requires that the binding site of the regulatory protein at least partially overlaps the regulated promoter [9,10].

Modulating the level of DNA supercoiling has also been proposed as a possible mechanism to regulate gene expression at a distance [11,12]. It is well known that expression of a number of bacterial genes is sensitive to the level of supercoiling [13–17]. Perturbation of DNA supercoiling in vivo by environmental factors, mutations in the genes encoding DNA topoisomerases or certain antibiotics influences the expression of particular genes [18–22]. These factors, however, affect gene expression very likely through a global effect on DNA supercoiling rather than acting specifically and locally [23]. It has been further proposed that a structural change at a specific site of DNA may modulate the level of local supercoiling, which in turn affects an event at another site in the same topological domain [24]. The question is, therefore, whether binding of a regulatory protein to an operator site is able to accomplish the structural change which can influence the level

*Corresponding author. Fax: (64) (6) 351 8032.
E-mail: agpfxm@pnv.palm.cri.nz

Abbreviations: bp, basepair(s); RNA Pol, RNA polymerase; pur, purine; pyr, pyrimidine.

SSDI S0014-5793(96)00139-1

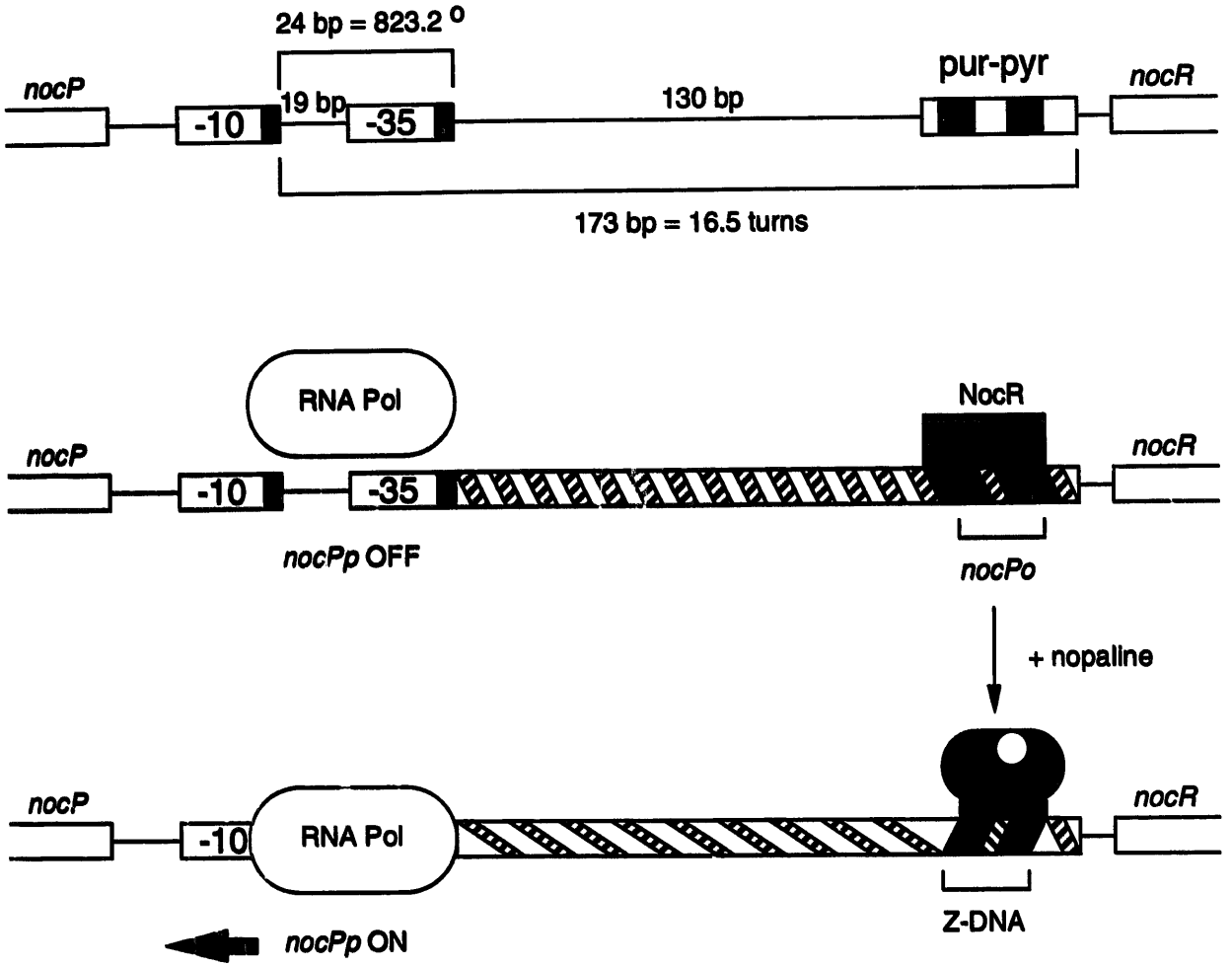

Fig. 1. Architecture of the *noc* promoter region and proposed model for regulation of the *nocP* promoter. The *top part* exhibits the structure of the *noc* promoter region. The −10 and −35 hexamers of the *nocP* promoter are represented by *boxes*, and the first T residues of the hexamers separated by 24 bp are labelled by *shadowing*. The 18 bp alternating purine–pyrimidine sequence 130 bp upstream of *nocPp* is labelled by *pur–pyr*. Inside this region, the tandem palindromes of the operator (*nocPo*) are represented by *filled boxes*. The coding portion of the *nocP* and *nocR* genes are labelled by *brackets*. The *middle part* shows the repressed state of the *nocP* promoter in the absence of nopaline. In this state the NocR protein binds to the *nocP* operator which is in B-configuration, and the *nocP* promoter is turned off, because its −10 and −35 hexamers are not positioned optimally for the transcription. The *bottom part* demonstrates the activation of the promoter. Nopaline (*small open circle*) binds to the NocR protein and alters its conformation. It is proposed that the conformationally altered NocR protein mediates the transition all or part of the operator into a Z-configuration. This B to Z conversion removes helical turns and supercoils from the downstream region, resulting in the correct spatial arrangement of the hexamers of the *nocP* promoter, which allows transcription. For more detailed explanation of the model, please refer to the text. The helical turns of the DNA double-helix are represented by *striped diagonal boxes*. The number of the turns shown in the figure does not reflect the actual number of the turns in the region (see text). For clarity, supercoils of the DNA are not shown.

of local supercoiling between the operator and the promoter sites to repress or activate transcription of the promoter.

It is well known that synthetic, alternating purine–pyrimidine sequences are able to form both B and Z-DNA under certain conditions [25], and that transition of the DNA double helix between the B and Z-forms affects the level of DNA supercoiling. Thus a B to Z transition leads to supercoil relaxation, while a Z to B transition introduces supercoils, respectively [26]. It was postulated that regulatory proteins may bind to Z-DNA and B–Z junctions [26,27], and may induce a B to Z or a Z to B transition in response to inducer molecules which bind to them and modulate their conformation [25]. Thus, regulatory protein-mediated B–Z transitions could influence DNA supercoiling, which in turn affects the spatial relationship between the −10 and −35 hexamers of a promoter, an essential component for effective transcription [28]. Consequently, gene expression could be regulated by this mechanism. In support of the above proposal, putative Z-DNA-forming sequences have been reported in the 5′ untranslated region of a number of genes [29], and in addition Z-DNA binding proteins have been isolated from different organisms [30–32]. Below, a bacterial operon is described, and a theoretical model presented which proposes that activation of a distant promoter may occur via induction a B to Z transition of a remote operator by a regulatory protein.

3. The structure of the *noc* operon of *Agrobacterium tumefaciens*

In *Agrobacterium tumefaciens* the *noc* (*no*paline catabolism)

genes encode proteins for the uptake and catabolism of nopaline, an amino acid-like molecule. Nopaline is produced by plant tumours induced by certain *A. tumefaciens* strains, and serves as a nutrient for the bacteria [33]. The uptake genes of the Ti plasmid pTiT37 form an operon and are regulated by the product of the *nocR* gene transcribed divergently from the uptake operon [34]. In the absence of nopaline, the promoter (*nocPp*) of the uptake genes is repressed, but in the presence of nopaline it is activated. The structure of the intervening region between the *nocP* and *nocR* genes is show in Fig. 1. Mobility shift assay, DNaseI footprinting, and mutagenesis experiments confirmed that the NocR regulatory protein binds to a single operator (*nocPo*) in both the absence and presence of nopaline, and also demonstrated that there is no operator site in or around the *nocP* promoter, whose position was revealed by functional and sequence analyses [34,35]. The distance between the *nocP* promoter and the operator is 131 bp, therefore the position of the operator, relative to the promoter, can be termed remote. Previous results also confirmed that there is no looping between the operator and the promoter, and that the NocR protein does not bend its target sequence [34,35]. It was proposed, therefore, that the expression of *nocPp* may be regulated by the NocR protein by modulating the supercoil level of the intervening sequence between the *nocP* operator and the *nocP* promoter [34,35].

Some previous results suggest the possibility that the *nocP* promoter is sensitive to changes in the level of DNA supercoiling. First, the −10 hexamer of the promoter has a CG-rich TCCCGT sequence, and the −10 and −35 hexamers are separated by a 19 bp spacer [36], which features have been reported for supercoiling-sensitive promoters [37,38]. Second, it has been demonstrated that the absence of a productive transcription divergent to *nocPp*, which decreases the level of the local supercoiling (i.e. relax DNA), increases expression of the *nocP* promoter [36]. Third, carbon starvation which relaxes DNA [39], and high osmolarity, which increases supercoiling [40], enhances and reduces the expression of the *nocP* promoter, respectively (Marincs, unpublished results).

4. Possible regulation of the *nocP* promoter

The expression of the *nocP* promoter is enhanced by conditions which lead to a decrease of DNA supercoiling (see above), and therefore it is possible that activation of the promoter may occur by a mechanism which involves decreasing the local level of DNA supercoiling; for example by the transition of a sequence from the B to the Z-form. The *noc* operator overlaps an 18 bp alternating purine–pyrimidine sequence (Fig. 1), which is the type of structure that may form a Z-DNA structure. Although there is no direct experimental evidence to date that the *noc* operator exists in both the B and Z-forms, there are some indications which intimate this possibility. For example, in the DNaseI footprints of the NocR protein–operator complex [35], characteristic hypersensitive sites were found, which have been reported to be present at B–Z junctions [41]. Furthermore, under conditions which favour the B to Z transition of DNA, the UV spectrum of the *noc* operator DNA displays certain changes (Marincs, unpublished results), which are known for Z-DNA [42]. Finally, the *noc* operator, inserted into a high copy number plasmid, causes relaxation of negative supercoils of the plasmid depending on the overall level of DNA supercoiling, which also indicates the possible formation of Z-DNA by the operator sequence [36]. Below we describe that even partial transition of this sequence from B to Z-form may turn on the *nocP* promoter.

The relative orientation of the −10 and −35 hexamers of σ^{70} promoters is an important factor for the RNA polymerase–promoter interaction [22]. Promoters with a 17 bp spacer between the −10 and −35 hexamers are optimally transcribed by the RNA polymerase, while deletion of basepair(s) from or insertion into the spacer region reduces promoter activity [37]. A 17 bp spacer means that the first T residues of the −35 and −10 hexamers are separated by 22 bp, and therefore the rotation between them is 754.6°, calculating with an average rotation of 34.3°/bp [43]. In contrast, the *nocP* promoter has a 19 bp spacer, consequently 24 bp and 823.2° between the two T residues (Fig. 1). This spatial conformation is suboptimal for the transcription of the promoter, and therefore the promoter is repressed. For activation of the promoter, the two T residues must be brought to the optimal angle (754.6°). This can be achieved by a −68.6° rotation of the first T residue of the −35 hexamer relative to the first T of the −10 hexamer.

The distance between the first T of the −10 hexamer of the *nocP* promoter and the end of the putative Z-DNA forming sequence positioned upstream is 173 bp (Fig. 1). The number of helical turns over this distance can be calculated using the equation

$$\Theta = (T \times 360) \div d \quad (1)$$

where: Θ is the average rotation between two neighbouring basepairs; T is the number of helical turns; 360 is the rotation per helical turn; and d is the length of the sequence in basepairs [44]. When the *nocP* promoter is repressed, then $\Theta = 34.3$, $d = 173$, and therefore $T \approx 16.5$ in the *noc* promoter region. If the −68.6° twist required for the activation of the promoter occurs, this would result in a 754.6° rotation between the two T residues, giving an average angle of 31.4°/bp. Assuming that the same twist occurs in the upstream region as well, the number of helical turns of the 173 bp region in the activated state using (1) is:

$$31.44 = (T \times 360) \div 173$$

giving a value of 15.1 for T. Consequently, the removal of $16.5 - 15.1 = 1.4$ helical turns is needed to activate the *nocP* promoter.

When n basepairs undergo a B to Z-form transition, this results in the removal of helical turns [45], the extent of which can be calculated from the equation

$$(n \div 10.5) + (n \div 12) = \Delta T \quad (2)$$

where n is the number of basepairs involved in the transition, while 10.5 and 12 are the number of basepairs per helical turn in B and Z-DNA, respectively, and ΔT is the number of the removed helical turns. From the calculations above, activation of the *nocP* promoter requires removal of 1.4 helical turns. Substituting this value into equation (2)

$$(n \div 10.5) + (n \div 12) = 1.4$$

which gives $n \approx 8$. This means that flipping of eight basepairs of the purine–pyrimidine sequence overlapping the *noc* operator is sufficient to remove 1.4 helical turns to activate the *nocP* promoter.

5. The model

The following model is proposed for the regulation of the *nocP* promoter (Fig. 1). In the absence of nopaline, the NocR protein binds to the *noc* operator which is in a B-DNA configuration. In this stage, the rotation between the first T residues of the −35 and −10 hexamers of the *nocP* promoter is suboptimal for efficient transcription, thus the promoter is repressed. When nopaline is present, its binding to the NocR regulatory protein induces a change in the conformation of the protein. As a consequence, the *noc* operator turns, at least partially, into a Z-configuration, which removes helical turns from the intervening region between the operator and the promoter. Therefore, the −35 hexamer of the *nocP* promoter rotates to the appropriate angle relative to the −10 hexamer which is optimal for transcription, and thus the promoter becomes activated.

This model implies that removal of helical turns activates the promoter. However, the *nocP* promoter has been found to be transcribed more efficiently when DNA is less supercoiled, i.e more relaxed. Furthermore, formation of Z-DNA is also known to remove supercoils from the DNA [26]. One might ask, therefore, how the removal of the helical turns and DNA supercoiling are connected in this case.

DNA exists in a negatively supercoiled form inside the cell, and can be characterised by the linking number (L), which can be expressed by the equation

$$L = T + W \tag{3}$$

where T and W are the numbers of helical turns and supercoils, respectively [46]. In a negatively supercoiled DNA, W is negative, thus

$$L = T + (-W). \tag{4}$$

Although the particular number of negative supercoils in the *noc* promoter region is not known, the linking number in the repressed state can be described with the equation

$$L_R = 16.5 + (-W_R) \tag{5}$$

where 16.5 is the number of the helical turns (see above), L_R and $-W_R$ are the linking number and the number of negative supercoils of the *noc* promoter region, respectively. When the promoter is activated, then T=15.1 and (4) becomes

$$L_A = 15.1 + (-W_A) \tag{6}$$

where L_A and $-W_A$ are the linking number and the numbers of negative supercoils in the activated state, respectively. Since the linking number must be constant [46], thus

$$L_R = L_A \tag{7}$$

and therefore from (5) and (6)

$$16.5 + (-W_R) = 15.1 + (-W_A). \tag{8}$$

Rearranging (8)

$$-W_R + 1.4 = -W_A$$

therefore

$$-W_R < -W_A$$

and

$$[-W_R] > [-W_A].$$

This means that the numbers of negative supercoils are higher in the repressed than in the activated state, i.e. during the activation the *noc* promoter region becomes less supercoiled, i.e. more relaxed. Thus, removal of supercoils and helical turns occurs simultaneously when the *nocP* promoter is activated, and therefore it is perspicuous now that factors and conditions which perturb DNA supercoiling influence the expression of the *nocP* promoter.

6. Conclusions

In this paper, we have presented a model which implies a possible molecular mechanism for the regulation of a promoter at a distance. In this model, a regulatory protein mediates a B to Z transition of a remote operator in response to an inducer molecule, and this transition causes the removal of both helical turns and negative supercoils from the intervening region between the operator and the promoter. This conformational changes bring the −10 and −35 hexamers of the regulated promoter to an optimal spatial orientation and therefore the promoter becomes activated.

Previously, a similar model has been presented by Sinden, proposing that transitions between the B and Z-forms of DNA may turn promoters on and off [47]. In contrast to Sinden's hypothesis, our model describes that which factors can mediate a B–Z transition, and also explains that what can be the biological role of an overlapping regulatory element (an operator) and a putative Z-DNA forming sequence.

Several lines of indirect evidence led us to develop the model, however direct evidence that the *noc* operator can exist in both the B and Z-form is not yet available. Although methodologies exist for investigating B-DNA in vivo, these techniques have been developed for synthetic purine–pyrimidine sequences and for *Escherichia coli* plasmids, and therefore they are not directly applicable to our *Agrobacterium* system. Furthermore, since overlapping sequences are involved in the process, simultaneous structural and functional analyses of these regions would be demanding. The model presented in this paper is however supported by previously described theoretical conceptions of others, and we are now in the position to use it as a working hypothesis for further investigations that gene expression may be regulated through modulation of the supercoiling level of DNA involving putative Z-DNA forming sequences and regulatory proteins binding to them.

Acknowledgements: We thank Drs Richard Biggs and Paul Reynolds for their critical comments on the manuscript.

References

[1] Collado-Vides, J., Magasnik, B. and Gralla, J.D. (1991) Microbiol. Rev. 55, 371–394.
[2] Ishihama, A. (1993) J. Bacteriol. 175, 2483–2489.
[3] Schleif, R. (1992) Annu. Rev. Biochem. 61, 199–223.
[4] Matthews, K.S. (1992) Microbiol. Rev. 56, 123–136.
[5] Gralla, J.D. (1989) Cell 57, 193–195.
[6] Su, W., Porter, S., Kustu, S. and Echols, H. (1990) Proc. Natl. Acad. Sci. USA 87, 5504–5508.
[7] Pérez-Martín, J. and De Lorenzo, V. (1995) J. Bacteriol. 177, 3758–3763.
[8] Adhya, S. and Garges, S. (1990) J. Biol. Chem. 265, 10797–10800.
[9] van der Vliet, P.C. and Verrijzer, C.P. (1993) BioEssays 15, 25–32.

[10] Pérez-Martín, J., Rojo, F. and De Lorenzo, V. (1994) Microbiol. Rev. 58, 269–290.
[11] Smith, G.R. (1981) Cell 24, 599–600.
[12] Pruss, G.J. and Drlica, K. (1989) Cell 56, 521–523.
[13] Whitehall, S., Austin, S. and Dixon, R. (1992) J. Mol. Biol. 225, 591–607.
[14] Scarlato, V., Aricò, B. and Rappuoli, R. (1993) J. Bacteriol. 175, 4764–4771.
[15] Grau, R., Gardiol, D., Glikin, G.C. and de Mendoza, D. (1994) Mol. Microbiol. 11, 933–941.
[16] Sun, L. and Fuchs, J.A. (1994) J. Bacteriol. 176, 4617–4626
[17] Wu, Y. and Datta, P. (1995) Mol. Gen. Genet. 247, 764–767.
[18] Sanzey, B. (1979) J. Bacteriol. 138, 40–47.
[19] Jovanovich, S.B. and Lebowitz, J. (1987) J. Bacteriol. 169, 4431–4435.
[20] Higgins, C.F., Dorman, C.J., Stirling, D.A., Waddell, L., Booth, J.R., May, G. and Bremer, E. (1988) Cell 52, 569–584.
[21] Drlica, K. (1992) Mol. Microbiol. 6, 425–433.
[22] Wang, J-Y. and Syvanen, M. (1992) Mol. Microbiol. 6, 1861–1866.
[23] Dorman, C.J. (1995) Microbiology 141, 1271–1280.
[24] Wang, J.C. and Giaever, G.N. (1988) Science 240, 300–304.
[25] Rich, A., Nordheim, A. and Wang, A. H-J. (1984) Annu. Rev. Biochem. 53, 791–846.
[26] Stirdivant, S.M., Klysik, J. and Wells, R.D. (1982) J. Biol. Chem. 257, 10159–10165.
[27] Davis, D.R. and Zimmerman, S. (1980) Nature 283, 11–12.
[28] Jyothirmai, G. and Mishra, R.K. (1994) FEBS Lett. 340, 189–192.
[29] Wells, R.D., Collier, D.A., Hanvey, J.C., Shimizu, M. and Wohlrab, F. (1988) FASEB J. 2, 2939–2949.
[30] Lafer, E.M., Sousa, R., Rosen, B., Hsu, A. and Rich, A. (1985) Biochemistry 24, 5070–5076.
[31] Lafer, E.M., Sousa, R.J. and Rich, A. (1988) J. Mol. Biol. 203, 511–516.
[32] Herbert, A.G., Spitzner, J.R., Lowenhaupt, K. and Rich, A. (1993) Proc. Natl. Acad. Sci. USA 90, 3339–3342.
[33] Dessaux, Y., Petit, A. and Tempé, J. (1992) in: Molecular Signals in Plant–Microbe Communications (Verma, D.P.S. ed.) CRC Press, Boca Raton, FL, pp. 109–136.
[34] Marincs, F. and White, D.W.R. (1994) Mol. Gen. Genet. 244, 367–373.
[35] Marincs, F. and White, D.W.R. (1993) Mol. Gen. Genet. 241, 65–72.
[36] Marincs, F. and White, D.W.R. (1995) J. Biol. Chem. 270, 12339–12342.
[37] Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319–353.
[38] Condee, C.W. and Summers, A.O. (1992) J. Bacteriol. 174, 8094–8101.
[39] Balke, V.L. and Gralla, J.D. (1987) J. Bacteriol. 169, 4499–4506.
[40] Hsieh, L-S., Rouviere, J. and Drlica, K. (1991) J. Bacteriol. 173, 3914–3917.
[41] Winkle, S.A., Aloyo, M.C., Lee-Chee, T., Morales, N., Zambrano, T.Y. and Sheardy, R.D. (1992) J. Biomol. Struct. Dyn. 10, 389–402.
[42] Jovin, T.M., Soumpasis, D.M. and McIntosh, L.P. (1987) Annu. Rev. Phys. Chem. 38, 521–560.
[43] Sinden, R.R. (1994a) DNA structure and function. Academic Press, San Diego, CA, p. 23.
[44] Sinden, R.R. (1994b) DNA structure and function. Academic Press, San Diego, CA, p.103.
[45] Sinden, R.R. (1994c) DNA structure and function. Academic Press, San Diego, CA, p. 190.
[46] Sinden, R.R. (1994d) DNA structure and function. Academic Press, San Diego, CA, p. 99.
[47] Sinden, R.R. (1994e) DNA structure and function. Academic Press, San Diego, CA, pp. 207–213.